Rank Revealing QR Factorizations

Benjamin Daniel   Arvind Saibaba   Ilse Ipsen

NC State University

July 24, 2020

Overview

- Motivation and Definitions
- Literature
- Our Contributions
  - Theorems
  - Numerical Results

Rank Revealing Factorizations

Finding the numerical rank of a matrix has applications in subset selection, least squares, regularization, matrix approximation, etc (Chan and Hansen 1991).

The SVD is the "best" rank-revealing factorization. However, it is computationally expensive, and does not have interpretability.

Many researchers have used QR factorizations to determine numerical rank.

Rank-Revealing QR Factorization

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $m > n$ and $1 \leq k < n$. A Rank-Revealing QR of $\mathbf{A}$ is a QR factorization of $\mathbf{A\Pi}$,

$$\mathbf{A} \begin{bmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 \end{bmatrix} = \mathbf{Q} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22} \end{bmatrix}$$

where $\mathbf{R}_{11} \in \mathbb{R}^{k \times k}$ and $\mathbf{\Pi}$ is a permutation chosen such that there are polynomials $p_1(n,k)$ and $p_2(n,k)$ such that for $1 \leq i \leq k$ and $1 \leq j \leq n-k$, we have

$$\frac{\sigma_i(\mathbf{R}_{11})}{p_1(n,k)} \leq \sigma_k(\mathbf{A}) \qquad \text{and} \qquad \sigma_j(\mathbf{R}_{22}) \leq \sigma_{j+k}(\mathbf{A}) p_2(n,k).$$

The Rank-Revealing QR Factorization in Literature

- Businger and Golub (1965)
- Golub, Klema, and Stewart (1976)
- Hong and Pan (1992)
- Chandrasekaran and Ipsen (1994)
- Gu and Eisenstat (1996)

The Rank-Revealing QR Factorization in Literature

- Businger and Golub (1965)
- Golub, Klema, and Stewart (1976)
- Hong and Pan (1992)
- Chandrasekaran and Ipsen (1994)
- Gu and Eisenstat (1996)

Golub, Klema, and Stewart's Algorithm

Key Observation: $\mathbf{A\Pi} = \mathbf{U\Sigma V}^{\mathrm{T}}\mathbf{\Pi}$, so permuting columns of $\mathbf{A}$ is equivalent to permuting columns of $\mathbf{V}^{\mathrm{T}}$.

Let

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_k & \mathbf{V}_\perp \end{bmatrix}.$$

Instead of choosing $\mathbf{\Pi}$ to select columns of $\mathbf{A}$, choose $\mathbf{\Pi}$ such that $\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1$ is large, ($\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1 \approx \mathbf{I}_k$).

This $\mathbf{\Pi}$ should also select columns of $\mathbf{A}$ such that $\mathbf{A\Pi}_1$ ($\mathbf{R}_{11}$) is well-conditioned.

Bounds on Singular Values

Theorem: Based on Golub, Klema Stewart (1976)

Given an permutation $\mathbf{\Pi}$ such that $\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1$ is nonsingular, we have the following bounds:

- $\sigma_j(\mathbf{R}_{22}) \leq \sigma_{k+j}(\mathbf{A})\|(\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2 \qquad 1 \leq j \leq n-k.$
- $\frac{\sigma_i(\mathbf{A})}{\|(\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} \leq \sigma_i(\mathbf{R}_{11}) \qquad 1 \leq i \leq k.$

Consequence of Interlacing Property of Singular Values

- $\sigma_i(\mathbf{R}_{11}) \leq \sigma_i(\mathbf{A})$ for $1 \leq i \leq k$.
- $\sigma_{k+j}(\mathbf{A}) \leq \sigma_j(\mathbf{R}_{22})$ for $1 \leq j \leq n-k$.

Bounds on Singular Values

We denote by $0 \leq \theta_1 \leq \ldots \leq \theta_k \leq \frac{\pi}{2}$, the principal angles between $\mathcal{R}(\mathbf{V}_k)$ and $\mathcal{R}(\mathbf{\Pi}_1)$.

Then $\frac{1}{\|(\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} = \cos(\theta_k)$. Using the results from the previous slide:

- $\sigma_{k+j}(\mathbf{A}) \leq \sigma_j(\mathbf{R}_{22}) \leq \frac{\sigma_{k+j}(\mathbf{A})}{\cos(\theta_k)} \qquad 1 \leq j \leq n-k.$
- $\sigma_i(\mathbf{A})\cos(\theta_k) \leq \sigma_i(\mathbf{R}_{11}) \leq \sigma_i(\mathbf{A}) \qquad 1 \leq i \leq k.$

Strong Rank-Revealing QR (Gu and Eisenstat 1996)

A rank-revealing QR is strong if for some parameter $f \geq 1$, we have the following bounds of the singular values of $\mathbf{R}_{11}$ and $\mathbf{R}_{22}$ as follows

$$\frac{\sigma_i(\mathbf{A})}{\sqrt{1 + f^2 k(n-k)}} \leq \sigma_i(\mathbf{R}_{11}) \quad 1 \leq i \leq k$$

$$\sigma_j(\mathbf{R}_{22}) \leq \sqrt{1 + f^2 k(n-k)}\sigma_{j+k}(\mathbf{A}) \quad 1 \leq j \leq n-k.$$

In addition, $\mathbf{R}_{11}$ nonsingular and we can bound the elements of $\mathbf{R}_{11}^{-1}\mathbf{R}_{12}$ in magnitude by

$$|(\mathbf{R}_{11}^{-1}\mathbf{R}_{12})_{ij}| \leq f, \quad 1 \leq i \leq k,\ 1 \leq j \leq n-k.$$

Our Approach

- Based on Golub, Klema, and Stewart: Reveal the rank of $\mathbf{A}$ by doing a strong rank-revealing QR on $\mathbf{V}_k^{\mathrm{T}}$.
- $\mathbf{V}_k$ is expensive to compute.
- Use randomized SVD to find a $\mathbf{W}$ with orthonormal columns such that $\mathbf{W} \approx \mathbf{V}_k$.
- Perform strong rank-revealing QR on $\mathbf{W}^{\mathrm{T}}$ instead of $\mathbf{V}_k$.

Our Approach

This leads to the following approach:

Algorithm 1: Our Approach to Rank Revealing QR

Use randomization techniques to find $\mathbf{W} \approx \mathbf{V}_k$.
Find strong RRQR decomposition of $\mathbf{W}^{\mathrm{T}}$, $\mathbf{W}^{\mathrm{T}}\mathbf{\Pi} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$.
Compute QR decomposition of $\mathbf{A\Pi}$.

Generalized $\mathbf{R}_{22}$ Bound

Recall from earlier:

The $\mathbf{R}_{22}$ Bound

If $\mathbf{V}_k^{\mathrm{T}}\boldsymbol{\Pi}_1$ is nonsingular, then

$$\sigma_j(\mathbf{R}_{22}) \leq \sigma_{k+j}(\mathbf{A})\|(\mathbf{V}_k^{\mathrm{T}}\boldsymbol{\Pi}_1)^{-1}\|_2 \quad 1 \leq j \leq n-k.$$

Can we generalize this from $\mathbf{V}_k$ to $\mathbf{W}$?

Theorem: Generalized $\mathbf{R}_{22}$ Bound

Let $\mathbf{W} \in \mathbb{R}^{n\times k}$ have orthonormal columns with $\mathbf{W}^{\mathrm{T}}\boldsymbol{\Pi}_1$ nonsingular. Then

$$\sigma_j(\mathbf{R}_{22}) \leq \sigma_j(\mathbf{A}(\mathbf{I}-\mathbf{W}\mathbf{W}^{\mathrm{T}}))\|(\mathbf{W}^{\mathrm{T}}\boldsymbol{\Pi}_1)^{-1}\|_2 \qquad \text{for } 1 \leq j \leq n-k.$$

Generalized $\mathbf{R}_{22}$ Bound

Theorem: Generalized $\mathbf{R}_{22}$ Bound

Let $\mathbf{W} \in \mathbb{R}^{n\times k}$ have orthonormal columns with $\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1$ nonsingular. Then

$$\sigma_j(\mathbf{R}_{22}) \leq \sigma_j(\mathbf{A}(\mathbf{I}-\mathbf{W}\mathbf{W}^{\mathrm{T}}))\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2 \qquad \text{for } 1 \leq j \leq n-k.$$

Let $\mathbf{W} = \mathbf{V}_k$ then

$$\sigma_j(\mathbf{A}(\mathbf{I}-\mathbf{W}\mathbf{W}^{\mathrm{T}})) = \sigma_j(\mathbf{A}(\mathbf{I}-\mathbf{V}_k\mathbf{V}_k^{\mathrm{T}})) = \sigma_{j+k}(\mathbf{A})$$

and the above statement becomes

$$\sigma_j(\mathbf{R}_{22}) \leq \sigma_{k+j}(\mathbf{A})\|(\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2.$$

Generalized $\mathbf{R}_{11}$ Bound

The $\mathbf{R}_{11}$ bound

If $\mathbf{V}_k^{\mathrm{T}}\boldsymbol{\Pi}_1$ is nonsingular, then

$$\frac{\sigma_i(\mathbf{A})}{\|(\mathbf{V}_k^{\mathrm{T}}\boldsymbol{\Pi}_1)^{-1}\|_2} \leq \sigma_i(\mathbf{R}_{11}) \qquad 1 \leq i \leq k.$$

Again, we wish to generalize this from $\mathbf{V}_k$ to $\mathbf{W}$.

Conjecture: Generalized $\mathbf{R}_{11}$ Bound

Let $\mathbf{W} \in \mathbb{R}^{n\times k}$ have orthonormal columns with $\mathbf{W}^{\mathrm{T}}\boldsymbol{\Pi}_1$ nonsingular. Then

$$\frac{\sigma_i(\mathbf{A}\mathbf{W}\mathbf{W}^{\mathrm{T}})}{\|(\mathbf{W}^{\mathrm{T}}\boldsymbol{\Pi}_1)^{-1}\|_2} \overset{?}{\leq} \sigma_i(\mathbf{R}_{11}) \qquad 1 \leq i \leq k.$$

$\mathbf{R}_{11}$ Bound Counterexample

The bound

$$\frac{\sigma_i(\mathbf{A}\mathbf{W}\mathbf{W}^{\mathrm{T}})}{\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} \leq \sigma_i(\mathbf{R}_{11}) \qquad 1 \leq i \leq k.$$

does NOT hold in general!

Example: Consider the case where $m = n = 2$ and $k = 1$ with

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \qquad \mathbf{W} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \text{and} \quad \mathbf{\Pi} = \mathbf{I}_2.$$

Then

$$\frac{\|\mathbf{A}\mathbf{W}\mathbf{W}^{\mathrm{T}}\|_2}{\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} = \frac{\sqrt{5}}{2} > 1 = \|\mathbf{R}_{11}\|_2.$$

$\mathbf{R}_{11}$ Bound Counterexample with Assumptions

Consider another case where $m = n = 2$ and $k = 1$ with

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

Then $\mathbf{A}$ has SVD $\mathbf{A} = \mathbf{U\Sigma V}^{\mathrm{T}}$, where

$$\mathbf{U} = \mathbf{I}_2, \quad \mathbf{\Sigma} = \begin{bmatrix} \sqrt{2} & \\ & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Let $0 < \epsilon < 1 - \frac{1}{\sqrt{2}}$ and choose $0 < \delta$ such that

$$\mathbf{W} = \begin{bmatrix} \frac{1}{\sqrt{2}} + \epsilon \\ \frac{1}{\sqrt{2}} - \delta \end{bmatrix}$$

has unit norm.

$\mathbf{R}_{11}$ Bound Counterexample with Assumptions

Consider again

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{V}_k = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \qquad \mathbf{W} = \begin{bmatrix} \frac{1}{\sqrt{2}} + \epsilon \\ \frac{1}{\sqrt{2}} - \delta \end{bmatrix}$$

Since the first element of $\mathbf{W}$ is larger, choose $\mathbf{\Pi} = \mathbf{I}$. Let $\mathbf{A\Pi} = \mathbf{QR}$, where $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = \mathbf{A}$. We compute

$$\frac{\|\mathbf{AWW}^{\mathrm{T}}\|_2}{\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} > 1 = \|\mathbf{R}_{11}\|_2.$$

We have

- $\mathbf{W}$ is an epsilon perturbation of $\mathbf{V}_k$.
- $\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1$ is optimal

The bound doesn't hold.

Generalized $\mathbf{R}_{11}$ Bound

Theorem: Generalized $\mathbf{R}_{11}$ Bound

Let $\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1$ be nonsingular and let

$$\sin\theta_k(\mathcal{R}(\mathbf{W}), \mathcal{R}(\mathbf{V_k})) < \frac{1}{\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2}.$$

Then

- $\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1$ is nonsingular.
- For $1 \leq i \leq k$,

$$\sigma_i(\mathbf{A})\left(\frac{1}{\|(\mathbf{W}^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} - \sin\theta_k(\mathcal{R}(\mathbf{W}), \mathcal{R}(\mathbf{V_k}))\right) \leq \sigma_i(\mathbf{R}_{11}).$$

If $\mathbf{W} = \mathbf{V}_k$, then $\sin\theta_k(\mathcal{R}(\mathbf{W}), \mathcal{R}(\mathbf{V_k})) = 0$ and the bound becomes

$$\frac{\sigma_i(\mathbf{A})}{\|(\mathbf{V}_k^{\mathrm{T}}\mathbf{\Pi}_1)^{-1}\|_2} \leq \sigma_i(\mathbf{R}_{11}).$$

Numerical Results

For our algorithm, we use a randomized range finder based on Algorithm 4.4 in Halko, Martinsson, and Tropp (2009) to find an orthonormal matrix $\mathbf{Q}$ with $\mathcal{R}(\mathbf{Q}) \approx \mathcal{R}(\mathbf{A})$

The right k dominant singular vectors of $\mathbf{AQQ}^{\mathrm{T}}$ gives us our desired $\mathbf{W}$.

Algorithm 2: Rank Revealing QR with Randomized SVD

Use Randomized SVD to find $\mathbf{W} \approx \mathbf{V}_k$.
Find strong RRQR decomposition of $\mathbf{W}^{\mathrm{T}}$, $\mathbf{W}^{\mathrm{T}}\mathbf{\Pi} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$.
Compute QR decomposition of $\mathbf{A\Pi}$.

Test matrices:

- Kahan Matrix
- Gravity Matrix

Numerical Results - $\mathbf{R}_{11}$

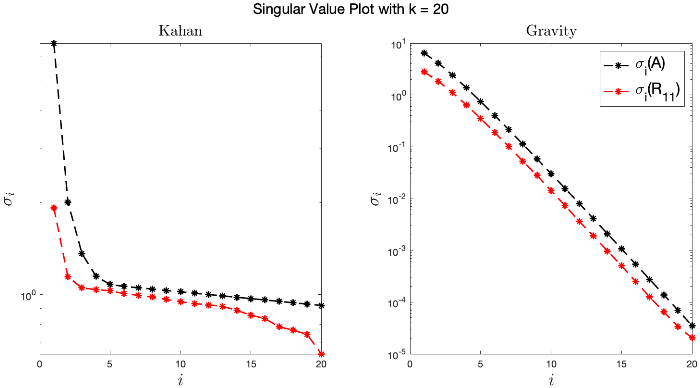

Figure: Singular Values of $\mathbf{A}$ and $\mathbf{R}_{11}$ for two test matrices.

Numerical Results - $\mathbf{R}_{11}$

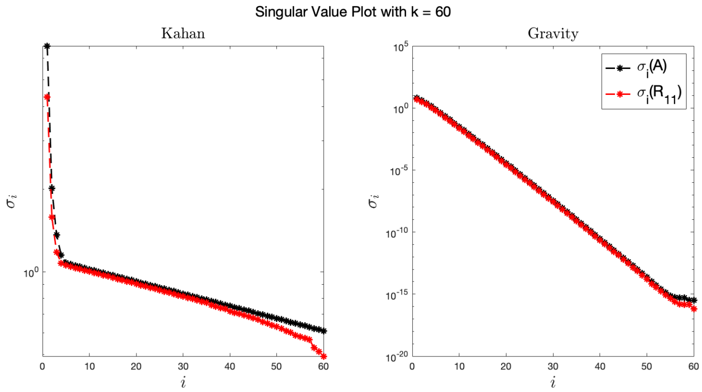

Figure: Singular Values of $\mathbf{A}$ and $\mathbf{R}_{11}$ for two test matrices.

Numerical Results - $\mathbf{R}_{22}$

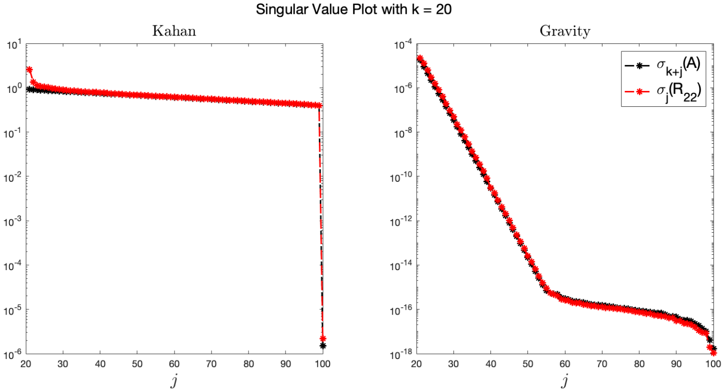

Figure: Singular Values of $\mathbf{A}$ and $\mathbf{R}_{22}$ for two test matrices.

Numerical Results - $\mathbf{R}_{22}$

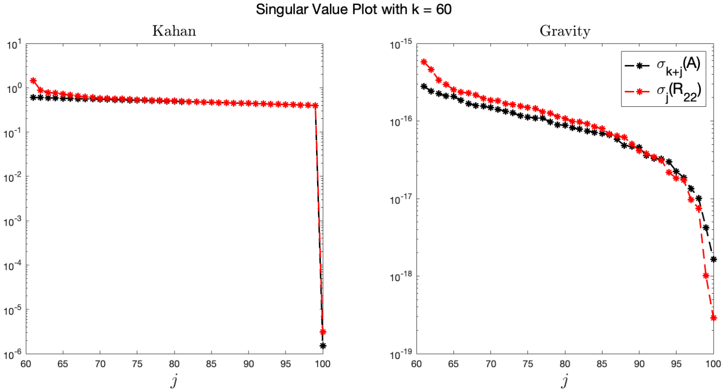

Figure: Singular Values of $\mathbf{A}$ and $\mathbf{R}_{11}$ for two test matrices.

Future Work

- Further explore lower bounds for $\mathbf{R}_{11}$ with minimal assumptions on $\mathbf{W}$.
- Construct bounds specific to our Algorithm.
- Finish the analysis of the computational cost of our algorithm.

References I

- P. Businger and G. Golub, *Linear least squares solutions by householder transformations*, Numerische Mathematik, 7 (1966), pp. 269–276.
- T. F. Chan and P. C. Hansen, *Some applications of the rank revealing qr factorization*, SIAM Journal on Scientific and Statistical Computing, 13 (1992), pp. 727–741.
- S. Chandrasekaran and I. C. F. Ipsen, *On rank-revealing factorisations*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 592–622.

References II

- G. Golub, V. Klema, and G. Stewart, *Rank degeneracy and least squares problems*, Tech. Rep. STAN-CS-76-559, Computer Science Department, Stanford University, 1976.
- M. Gu and S. C. Eisenstat, *Efficient algorithms for computing a strong rank-revealing qr factorization*, SIAM Journal on Scientific Computing, 17 (1996), pp. 848–869.
- N. Halko, P. G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Review, 53 (2011), pp. 217–288.
- Y. Hong and C.-T. Pan, *Rank-revealing qr factorizations and the singular value decomposition*, Mathematics of Computation, 58 (1992), pp. 213–232.