Derivatives of Randomized Matrix Approximation Algorithms

Joonsoo Lee

December 14, 2024

1 Introduction

Large matrices arise in many fields of science and engineering, such as optimization [1] numerical analysis [6], and uncertainty quantification [5]. For such large matrices $A \in \mathbb{R}^{m\times n}$ with $m, n >> 0$, randomized algorithms [3] are widely used to compute approximations and decompositions. These methods are able to reduce computational costs while also providing theoretical bounds on their behavior and error.

While much work has been done to develop and analyze algorithms approximating matrices, the behavior of differentiating such methods is unknown. In many practical applications, such as machine learning, understanding the sensitivity and bounds of these algorithms under differentiation is critical for error analysis. This project aims to generate preliminary discussion and demonstrate initial findings and analysis about such gap.

We first give a summary of the findings in the formative paper [3], focusing on the randomized range finding algorithm. Following, we then derive analytical bounds for the behavior of such algorithms under differentiation. Finally, we demonstrate numerical tests several algorithms, showing their performance in capturing and bounding derivative errors, and discuss these implications.

1.1 Overview

In this section, we give a brief overview of the mathematical formulation of the problem. The algorithms in [3] are based on finding a subspace that captures most of the action of a matrix. In mathematical terms, given matrix A and error tolerance ϵ, we want an orthonormal matrix $Q \in \mathbb{R}^{n\times \ell}$ such that $\ell = \ell(\epsilon)$ is monotonic and

$$\|A - QQ'A\| \leq \epsilon. \tag{1}$$

Note that the relationship $\ell(\epsilon)$ is not completely specified yet and the norm $\|\cdot\|$ is commonly the Frobenius ($\|\cdot\|_F$) or spectral norm ($\|\cdot\|_2$).

The main proto-algorithm described in [3] is:

- Draw a random $n \times \ell$ random matrix Ω (e.g. random Gaussian matrix).
- Form the product $Y = A\Omega$.
- Construct a $m \times \ell$ matrix Q whose columns form an orthonormal basis for the range of Y (e.g. $QR(Y)$).

As mentioned previously, this algorithm is shown to have some error bound for $\|A - QQ'A\|$. Our goal is to study the behavior of the derivative of this algorithm:

$$\|dA - d(QQ'A)\|. \tag{2}$$

2 Error Bounds in Nondifferentiated Algorithm

In this section, we review the theorems and analysis to obtain the error bound for (**??**). This is useful to understand the similar techniques used in the differentiated case. Furthermore, a naive way to obtain a approximation on dA is to run the same algorithm on dA itself, although this is not practical in many cases.

The analysis contains two sections, a deterministic error bound, which bounds (1) in terms of the random matrix Ω, and a probabilistic analysis which bounds the tails and expectation of such Ω terms.

2.1 Deterministic Error Bound

We first summarize the deterministic error analysis. To start, suppose $A \in \mathbb{R}^{m \times n}$ has the SVD $A = U\Sigma V$. Partition the matrices

$$A = U \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \end{bmatrix}, \tag{3}$$

where $\Sigma_1 \in \mathbb{R}^{k \times k}$. Then, let $\Omega \in \mathbb{R}^{n \times \ell}$ with $\ell \geq k$ and define

$$\Omega_1 = V_1'\Omega, \Omega_2 = V_2'\Omega. \tag{4}$$

So, our sample matrix $Y = A\Omega$ can be rewritten as

$$Y = U \begin{bmatrix} \Sigma_1\Omega_1 \\ \Sigma_2\Omega_2 \end{bmatrix}. \tag{5}$$

We also define the orthogonal projector, which has many useful properties contained in section 8 of [3].

Definition 2.1. An orthogonal projector is a Hermitian matrix P that satisfies $P^2 = P$. Clearly, this identity implies $0 \preceq P \preceq I$. Note that an orthogonal projector is completely determined by it's range and for a given matrix M, P_M is the unique orthogonal projector with $\text{range}(P_M) = \text{range}(M)$.

Since $QQ' = P_Y$, the challenge is to obtain bounds for

$$\|A - QQ'A\| = \|(I - P_Y)A\|. \tag{6}$$

Theorem 2.2 (9.1 in [3]). *Let A be $m \times n$ with singular value decomposition $A = U\Sigma V'$, and fix $k \geq 0$. Define $Y = A\Omega$ for a test matrix Ω. Partition Σ and define Ω_1 and Ω_2 as above. If Ω_1 has full row rank,*

$$\|(I - P_Y)A\|^2 \leq \|\Sigma_2\|^2 + \|\Sigma_2\Omega_2\Omega_1^\dagger\|^2, \tag{7}$$

with the spectral or Frobenius norm and $\dagger$ as the Moore–Penrose inverse.

The proof of this theorem in [3] utilizes the properties of the orthogonal projector P_Y and the relationship between Y and A.

2.2 Average Error Bound

In this section we show the bounds for the expectation of (7). For simplicity, we assume that Ω are Gaussian matrices. Usefully, a standard Gaussian matrix is rotationally invariant ($U'GV$ Gaussian if U and V are orthonormal). The following lemmas are particularly useful for our bounds.

Proposition 2.3 (10.1 in [3]). *Given matrices S and T with standard Gaussian matrix G, then*

$$\left(\mathbb{E}\|SGT\|_F^2\right)^{1/2} = \|S\|_F\|T\|_F, \tag{8}$$

$$\mathbb{E}\|SGT\|_2 \leq \|S\|_2\|T\|_F + \|S\|_F\|T\|_2. \tag{9}$$

Proposition 2.4 (10.2 in [3]). *Given $k \times (k+p)$ Gaussian matrix G with $k, p \geq 2$, then*

$$\left(\mathbb{E}\|G^\dagger\|_F^2\right)^{1/2} = \sqrt{\frac{k}{p-1}}, \tag{10}$$

$$\mathbb{E}\|G^\dagger\|_2 \leq \frac{e\sqrt{k+p}}{p}. \tag{11}$$

Theorem 2.5 (10.5 in [3]). *If $A \in \mathbb{R}^{m\times n}$ with singular values $\sigma_1 \geq \sigma_2 \geq ....$ Choose a target rank $k \geq 2$ and oversampling parameter $p \geq 2$ with $k + p \leq \min\{m, n\}$. Draw a $\Omega \in \mathbb{R}^{n\times(k+p)}$ standard Gaussian matrix and construct the sample matrix $Y = A\Omega$. Then, the expected error*

$$\mathbb{E}\|(I - P_Y)A\|_F \leq \left(1 + \frac{k}{p-1}\right)^{1/2} \left(\Sigma_{j>k}\sigma_j^2\right)^{1/2}. \tag{12}$$

Note that the Eckart-Young theorem [2] shows that $(\Sigma_{j>k}\sigma_j^2)^{1/2}$ is the minimum Frobenius norm error when approximating A with a rank k matrix.

Theorem 2.6 (10.6 in [3])**.** *Under the hypothesis of Theorem* (2.5)*,*

$$\mathbb{E}\|(I - P_Y)A\|_2 \leq \left(1 + \sqrt{\frac{k}{p-1}}\right)\sigma_{k+1} + \frac{e\sqrt{k+p}}{p}(\Sigma_{j>k}\sigma_j^2)^{1/2}. \tag{13}$$

Note that this bound implies

$$E\|(I - P_Y)A\|_2 \leq \left[1 + \sqrt{\frac{k}{p-1}} + \frac{e\sqrt{k+p}}{p}\sqrt{\min\{m,n\} - k}\right]\sigma_{k+1}, \tag{14}$$

so the average spectral norm error lies in a polynomial factor of σ_{k+1}.

In fact, [4] has shown that the minimum spectral-norm error when approximating A with a rank k matrix is σ_{k+1}. The proofs for these theorems utilizes the properties of Gaussian matrices applied on (7).

2.3 Tail Error Bounds

In this section we show the bounds for the tails of (7) to ensure that the bounds in expectation are useful.

Theorem 2.7. *With the same hypothesis of Theorem* (2.5) *and* $p \geq 4$*, for all* $u, t \geq 1$*,*

$$\|(I - P_Y)A\|_F \leq \left(1 + t\sqrt{12k/p}\right)(\Sigma_{j>k}\sigma_j^2)^{1/2} + ut\frac{e\sqrt{k+p}}{p+1}\sigma_{k+1}, \tag{15}$$

with failure probability at most $5t^{-p} + 2e^{-u^2/2}$*.*

The first term is the average error bound and the second term represents the deviation above the mean.

Theorem 2.8. *Under the same hypothesis of Theorem* (2.7)*,*

$$\|(I - P_Y)A\| \leq \left[\left(1 + t\sqrt{12k/p}\right)\sigma_{k+1} + t\frac{e\sqrt{k+p}}{p+1}(\Sigma_{j>k}\sigma_j^2)^{1/2}\right] + ut\frac{e\sqrt{k+p}}{p+1}\sigma_{k+1}, \tag{16}$$

with failure probability at most $5t^{-p} + 2e^{-u^2/2}$*.*

The proofs of these theorems again utilizes further properties of Gaussian matrices which can be seen in further detail in [3].

3 Error Bounds of Derivatives

Now, we will study the error of the differentiated algorithm (2). For convenience to the reader, we repeat the error we seek to bound:

$$\|dA - d(QQ'A)\|. \tag{17}$$

For simplicity, we assume $A \in \mathbb{R}^{n\times n}$, $A = A'$, and $A \succ 0$. This is can be common for many hessians $\nabla^2 F$ that arise in optimization.

3.1 Deterministic Error Bound

Following the same approach as [3], we first attempt to bound

$$\|dA - d(QQ'A)\| = \|(I - QQ')dA - d(QQ')A\| \leq \|(I - QQ')dA\| + \|d(QQ')A\|. \tag{18}$$

Lemma 3.1. *Suppose $QR = QR(A\Omega)$, $A \in \mathbb{R}^{n\times n}$, $\Omega \in \mathbb{R}^{n\times \ell}$, $n \geq \ell$. If $rank(A\Omega) = \ell$, then*

$$\|Q\| \leq \sqrt{\ell}, \tag{19}$$

$$\frac{1}{\|Q\|}\|A\Omega\| \leq \|R\|. \tag{20}$$

Proof. By definition of the QR decomposition, $Q'Q = I$, so

$$\|Q\| = \|I\| \tag{21}$$

for both the spectral and Frobenius norm due to unitary invariance. Clearly,

$$\|Q\|_F = \sqrt{\ell}, \|Q\|_2 = 1. \tag{22}$$

Furthermore, since $QR = A\Omega$, we have that

$$\|A\Omega\| \leq \|Q\|\|R\|, \tag{23}$$

from which the second inequality follows. □

Theorem 3.2. *Suppose $QR = QR(A\Omega)$, $A \in \mathbb{R}^{n\times n}$, $\Omega \in n \times \ell$, $n \geq \ell$. If $rank(A\Omega) = \ell$, then*

$$\|d(QQ')\| \leq \frac{2(\|Q\| + 3\|Q\|^2)}{\|A\Omega\|}\|dA\Omega\|. \tag{24}$$

Proof. Clearly,

$$\|d(QQ')\| \leq 2\|Q\|\|dQ\|. \tag{25}$$

We know

$$dQ = QX + dA\Omega R^{\dagger} - QQ'dA\Omega R^{\dagger}, \tag{26}$$

from appendix A. Since $R \in \mathbb{R}^{\ell \times \ell}$, we see that $Y = dA\Omega R^{\dagger}$ is equivalent to

$$RY = dA\Omega. \tag{27}$$

We can show

$$\|dA\Omega R^{\dagger}\| \leq \frac{\|dA\Omega\|}{\|R'\|} \leq \frac{\|dA\Omega\|\|Q\|}{\|A\Omega\|}. \tag{28}$$

The same approach can be used (along with the fact that $\|X\| \leq 2\|Q'dA\Omega R^{\dagger}\|$)

$$\|QX - QQ'dA\Omega R^{+}\| \leq \sqrt{\ell}K\frac{\|Q'dA\Omega\|}{\|R\|}. \tag{29}$$

The result is

$$\|d(QQ')\| \leq \frac{2(\|Q\| + 3\|Q\|^2)}{\|A\Omega\|}\|dA\Omega\|. \tag{30}$$

Note that since A is positive definite, we can bound

$$\frac{\|dA\Omega\|}{\|A\Omega\|} \leq \frac{\|dA\Omega\|}{\lambda_{min}(A)\|\Omega\|}. \tag{31}$$

□

The question remains to bound $\|(I - QQ')dA\|$. Although the form looks similar to the analysis in [3], the main difference is that the space of the projection $P_Y = QQ'$ may be completely different from the space of dA. It is unclear yet if any nontrivial bound can be obtained in this case. However, when we assume $dA = AdM$, for some dM, we can easily obtain the same bound as before from (7):

$$\|(I - QQ')dA\|^2 \leq (\|\Sigma_2\|^2 + \|\Sigma_2\Omega_2\Omega_1^{\dagger}\|^2)\|dM\|^2. \tag{32}$$

3.2 Average Error Bound

So, we now have the error bounds in the simple case:

$$\|dA - d(QQ'A)\| \leq \sqrt{(\|\Sigma_2\|^2 + \|\Sigma_2\Omega_2\Omega_1^{\dagger}\|^2)\|dM\|^2} + \frac{2(\|Q\| + 3\|Q\|^2)}{\|A\Omega\|}\|dA\Omega\|\|A\|. \tag{33}$$

However, this bound does not provide guarantees on convergence in expectation as $\ell \to n$. This is evident even with $dA = A$, as the second term would remain bounded by $\|A\|$ for all ℓ. Most likely, a separate analysis will need to be done without applying the triangle inequality to (18).

4 Numerical Tests

In this section we provide figures from numerical tests performed. We investigate the convergence of the norm error to 0 for several algorithms as $\ell \to n$. We compare three algorithms

- $Q = QR(dA\Omega)$ (a),
- algorithm we examined with $Q = QR(A\Omega)$ (b),
- $Q = QR(\Omega)$ (c).

Furthermore, our tests, we take $A = \tilde{\Omega}\tilde{\Omega}'$ where $\tilde{\Omega}$ is a 100×100 Gaussian matrix and 2500 samples of Ω.

Figure (1) shows this convergence for $dA = \varepsilon K$, with K as a Gaussian matrix. The blue line shows the error of algorithm (a), the red line shows the error of algorithm (b), and the green line shows the error of algorithm (c). The horizontal axis shows the dimension of the subspace ℓ ($\Omega \in \mathbb{R}^{n\times\ell}$), and the vertical axis is the average Frobenius norm error. Interestingly, algorithm (b) does not evern perform as well as algorithm (c) in this general case, which is a warning to excercise caution when differentiating the randomized algorithms in [3] without knowing the structure of dA. Furthermore, algorithm (b) also does not show monotonic convergence to 0. This could potentially be caused by a term of $\|A\Omega\|$ in the denominator from our analysis. Finally, algorithm (b) shows a very steep dropoff to 0 at the end. All of these signs point to there being very bad stability in outright differentiating the randomized algorithm in general cases.

Figure (2) shows convergence for $dA = \varepsilon AK$, which is the assumption we made to get the analytical error bound. The colors are for the same algorithms as in figure (2), and the axes are also the same. In this case, we can see that algorithm (b) outperforms algorithm (c) past a certain point, but in the beginning algorithm (c) performs better. Although this is a better sign, it is still worrying that algorithm (b) does not come close to the superlinear convergence of algorithm (a) even with the strict assumption on dA. Interstingly, algorithm (b) particularly performs poorly for low ℓ, which is suprising, as one would assume these algorithms converge quicker with low ℓ.

Both of these figures demonstrate that differentiating randomized matrix approximation algorithms is likely to be unstable, so careful care must be taken when doing so.

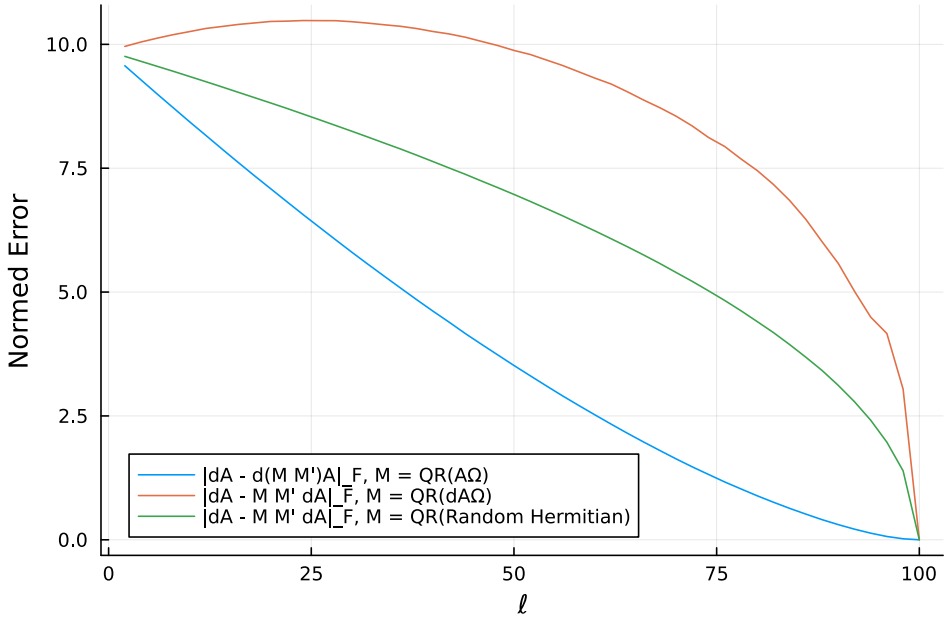

Figure 1: Finite Difference comparison for $n = 100$ and random Gaussian squared matrix A and dA.

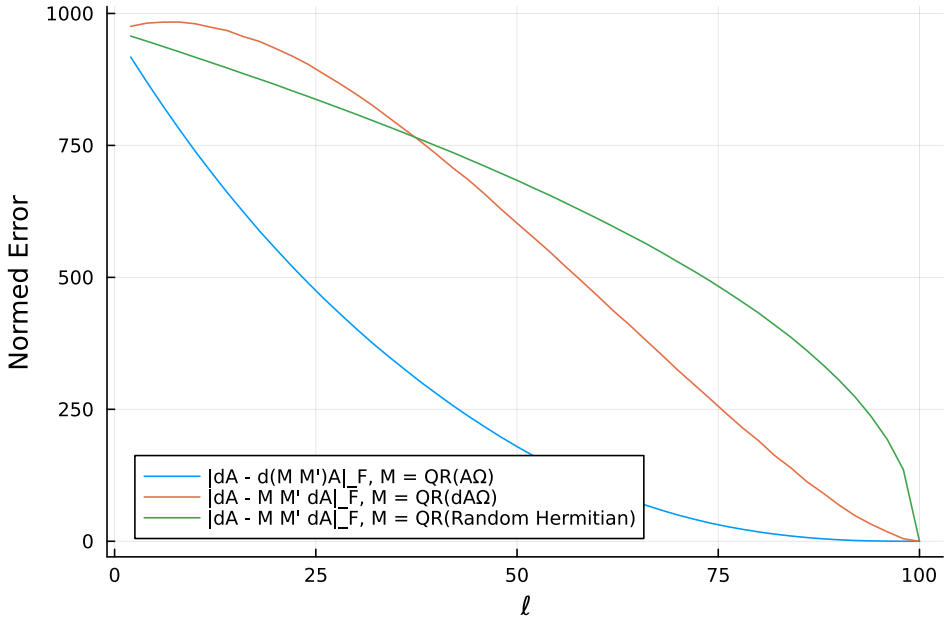

Figure 2: Finite Difference comparison for $n = 100$ and random Gaussian squared matrix A and dA.

5 Conclusion

In this paper we have summarized the theoretical work in [3] and analyzed the behavior of differentiating some algorithms in [3]. Mainly, we have outlined the challenges of applying a similar analysis to the differentiated algorithm and also shown numerical tests of poor behavior with respect to other methods.

There is much more work to be done. Firstly, analyzing other methods in [3], particularly power based methods which are shown to have better convergence and behavior would be a straightforward next step. Furthermore, obtaining a bound without using the triangle inequality in (18) may explain the behavior we see numerically. Finally, numerical tests on other types of scenarios could be insightful (e.g. non-Gaussian).

A Differentiating QR Decomposition [7]

In this section, we describe the exact form of the QR decomposition and its differentials to enforce uniqueness. We assume A is full column rank. Suppose we have

$$0 = A - QR, \tag{34}$$
$$0 = Q'Q, \tag{35}$$
$$0 = P_L \circ R \text{ (R upper triangular)}. \tag{36}$$

The differentials are easily calculated as

$$0 = dA - dQR - QdR, \tag{37}$$
$$0 = dQ'Q + Q'dQ, \tag{38}$$
$$0 = P_L \circ dR. \tag{39}$$

Defining $X = Q'dQ$, we can see

$$dR = Q'dA - XR. \tag{40}$$

From this, we can also calculate

$$dQ = QX + dAR^+ - QQ'dAR^+, \tag{41}$$

where $+$ is the Moore-Penrose pseudoinverse. Of course, we need to be able to calculate X without dQ. So, we see that

$$P_L \circ X = P_L \circ (Q'dAR^+), \tag{42}$$
$$X = (P_L \circ X) - (P_L \circ X)' \text{ (antisymmetric)}. \tag{43}$$

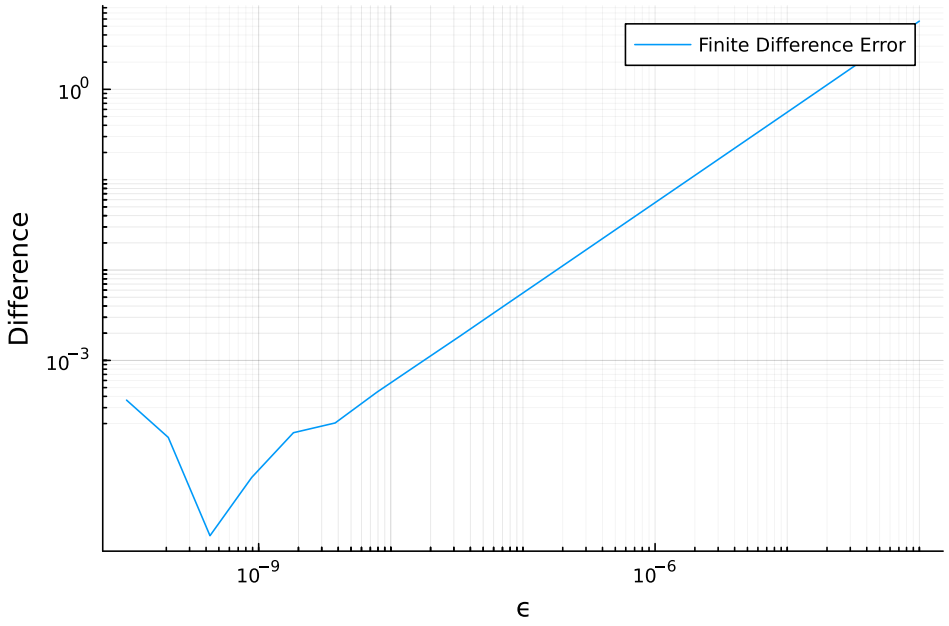

Figure 3: Finite Difference comparison for $n = 100$ and random Hermitian matrix A and dA.

In figure 3, we show the finite difference error of our explicit calculations. Mathematically, we compute

$$dQ_\epsilon = \frac{QR(A + \epsilon dA) - QR(A)}{\epsilon}, \tag{44}$$

and compare the difference

$$\left\| \frac{dQ}{\epsilon} - dQ_\epsilon \right\|. \tag{45}$$

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Mar. 2004.

[2] Carl Eckart and Gale Young. "The approximation of one matrix by another of lower rank". en. In: *Psychometrika* 1.3 (Sept. 1936), pp. 211–218.

[3] Nathan Halko, Per-Gunnar Martinsson, and Joel A Tropp. "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions". In: *SIAM review* 53.2 (2011), pp. 217–288.

[4] L Mirsky. "Symmetric gauge functions and unitarily invariant norms". In: *Q. J. Math.* 11.1 (1960), pp. 50–59.

[5] Timo Schorlepp et al. "Scalable methods for computing sharp extreme event probabilities in infinite-dimensional stochastic systems". In: *Statistics and Computing* 33.6 (Oct. 2023).

[6] Gilbert Strang and George Fix. *An Analysis of the Finite Element Methods, New Edition*. Philadelphia, PA: Wellesley-Cambridge Press, 2008.

[7] Sebastian F. Walter et al. "On evaluating higher-order derivatives of the QR decomposition of tall matrices with full column rank in forward and reverse mode algorithmic differentiation". In: *Optimization methods software.* 27.2 (2012-04), pp. 391–403.