A CLOSED-FORM UPDATE FOR ORTHOGONAL MATRIX DECOMPOSITIONS UNDER ARBITRARY RANK-ONE MODIFICATIONS

RALF ZIMMERMANN∗

Abstract. We consider rank-one adaptations \( X_{\text{new}} = X + ab^T \) of a given matrix \( X \in \mathbb{R}^{n \times p} \) with known matrix factorization \( X = UW \), where \( U \in \mathbb{R}^{n \times p} \) is column-orthogonal, i.e. \( U^T U = I \). Arguably the most important methods that produce such factorizations are the singular value decomposition (SVD), where \( X = UW = U \Sigma V^T \), and the QR-decomposition, where \( X = UW = QR \). By using a geometric approach, we derive a closed-form expression for a column-orthogonal matrix \( U_{\text{new}} \), whose columns span the same subspace as the columns of the rank-one modified \( X_{\text{new}} = X + ab^T \). This may be interpreted as a rank-one adaptation of the \( U \)-factor in the SVD or a rank-one adaptation of the \( Q \)-factor in the QR-decomposition, respectively. As a consequence, we obtain a decomposition for the adapted matrix \( X_{\text{new}} = U_{\text{new}} W_{\text{new}} \). In addition, the formula for \( U_{\text{new}} \) allows us to determine the subspace distance between the subspaces \( \text{ran}(X) = S \) and \( \text{ran}(X_{\text{new}}) = S_{\text{new}} \) without additional computational effort. In contrast to the existing approaches, the method does not require a numerical recomputation of the SVD or the QR-decomposition of an auxiliary matrix as an intermediate step. Rather, both \( U_{\text{new}} \) and \( W_{\text{new}} \) are obtained via elementary rank-one matrix updates in \( O(np) \) time for \( n > p \), which compares to \( O(np^2) \) for all other update methods known to the author.

Possible fields of applications include subspace estimation in computer vision, signal processing and adaptive model reduction.

Key words. singular value decomposition, rank-one update, subspace estimation, Grassmann manifold, rank-one subspace update, QR-decomposition

AMS subject classifications. 15B10, 15A83, 65F15, 65F25

1. Introduction. Investigations on the behavior of matrix decompositions under perturbations of restricted rank have a long tradition [5, 13, 6, 12, 11, 3, 2]. Of special importance in many applications are rank-one modifications \( X_{\text{new}} = X + ab^T \) of a given matrix \( X \in \mathbb{R}^{n \times p} \) with either known (thin) SVD \( X = UW = U \Sigma V^T \), where \( U \in \mathbb{R}^{n \times p} \), \( \Sigma, V \in \mathbb{R}^{p \times p} \), or known QR-decomposition \( X = QR \), where \( Q \in \mathbb{R}^{n \times p} \), \( R \in \mathbb{R}^{p \times p} \). In both cases, the matrix decomposition is of the form \( X = UW \) and the columns of \( U \) and \( Q \), respectively, provide an orthonormal basis for the range of \( X \), i.e., the subspace \( \text{ran}(X) \). For an introduction to subspace computations and updating matrix factorizations as well as additional references, the reader may consult [10] §6.4, 6.5.

The main original contribution of this note is a proof that an orthogonal matrix factor \( U_{\text{new}} \), i.e., an orthonormal basis of the updated column-span can be reached via a geodesic path on the Grassmann manifold [1], that starts in \( U \) (resp. \( Q \)) with a suitable tangent velocity \( \Delta \in \mathbb{R}^{n \times p} \), where \( \Delta \) is a rank-one matrix from the tangent space of the Grassmann manifold. This establishes a closed-form expression for \( U_{\text{new}} \), which, in turn, leads to a closed-form expression for the updated matrix factorization \( X_{\text{new}} = X + ab^T = U_{\text{new}} W_{\text{new}} \). It turns out that both \( U_{\text{new}} \) and \( W_{\text{new}} \) are obtained via a standard rank-one update on \( U \) and \( W \), respectively.

One may consider this as a formula for updating the orthonormal basis of a given subspace under arbitrary rank-one modifications.

∗Department of Mathematics and Computer Science, University of Southern Denmark (SDU) Odense, (zimmermann@imada.sdu.dk).
1.1. Notation and preliminaries. The \((p \times p)\)-identity matrix is denoted by \(I_p \in \mathbb{R}^{p \times p}\). The \((p \times p)\)-orthogonal group, i.e., the set of all square orthogonal matrices, is denoted by

\[
O_p = \{ R \in \mathbb{R}^{p \times p} | R^T R = R R^T = I_p \}.
\]

For a matrix \(X \in \mathbb{R}^{n \times p}\), the subspace spanned by the columns of \(X\) is called the range of \(X\) and is denoted by \(\mathcal{X} := \text{ran}(X) := \{ X \alpha \in \mathbb{R}^n | \alpha \in \mathbb{R}^p \} \subset \mathbb{R}^n\). We also speak of the subspace spanned by \(X\). The set of all \(p\)-dimensional subspaces \(\mathcal{X} \subset \mathbb{R}^n\) forms the Grassmann manifold

\[
\text{Gr}(n, p) := \{ \mathcal{X} \subset \mathbb{R}^n | \dim(\mathcal{X}) = p \}.
\]

The Stiefel manifold is the compact matrix manifold of all column-orthogonal rectangular matrices

\[
\text{St}(n, p) := \{ U \in \mathbb{R}^{n \times p} | U^T U = I_p \}.
\]

The Grassmann manifold can be realized as a quotient manifold of the Stiefel manifold

\[
\text{Gr}(n, p) = \text{St}(n, p)/O_p = \{ [U] | U \in \text{St}(n, p) \}, \tag{1.1}
\]

where \([U] = \{ UR | R \in O_p \}\) is the orbit, or equivalence class of \(U\) under actions of the orthogonal group. Hence, by definition, two matrices \(U, \hat{U} \in \text{St}(n, p)\) are in the same \(O_p\)-orbit if they differ by a \((p \times p)\)-orthogonal matrix:

\[
[U] = [\hat{U}] \iff \exists R \in O_p : U = \hat{U} R.
\]

A matrix \(U \in \text{St}(n, p)\) is called a matrix representative of a subspace \(\mathcal{U} \in \text{Gr}(n, p)\), if \(\mathcal{U} = \text{ran}(U)\). We will also consider the orbit \([U]\) and the subspace \(\mathcal{U} = \text{ran}(U)\) as the same object.

The tangent space \(T_{[U]} \text{Gr}(n, p)\) at a point \([U] \in \text{Gr}(n, p)\) can be thought of as the space of velocity vectors of differentiable curves on \(\text{Gr}(n, p)\) passing through \([U]\). For any matrix representative \(U \in \text{St}(n, p)\) of \([U] \in \text{Gr}(n, p)\), the tangent space of \(\text{Gr}(n, p)\) at \([U]\) is represented by

\[
T_{[U]} \text{Gr}(n, p) = \{ \Delta \in \mathbb{R}^{n \times p} | \Delta^T U = 0 \} \subset \mathbb{R}^{n \times p}, \tag{1.2}
\]

its canonical metric being \(\langle \Delta, \tilde{\Delta} \rangle_{Gr} = \text{tr}(\Delta^T \tilde{\Delta})\), \cite{8} §2.5. Endowing each tangent space with this metric turns \(\text{Gr}(n, p)\) into a Riemannian manifold. As in \cite{8}, we will make use throughout of the quotient representation \(\text{([1.1])}\) of the Grassmann manifold with matrices in \(\text{St}(n, p)\) acting as representatives in numerical computations.

Of special importance to this work are the geodesic lines on the Grassmann manifold. Geodesics on curved manifolds can be considered as the generalization of straight lines in flat, Euclidean spaces. From general differential geometry \cite{7}, it is known that a geodesic \(t \mapsto [U](t)\) is specified by a second-order differential equation and is thus uniquely determined by a starting point \([U] = [U](0)\) and a starting velocity \(\Delta = [\dot{U}](0) \in T_{[U]} \text{Gr}(n, p)\). This unique dependency gives rise to the so-called Riemannian exponential function

\[
\text{Exp}_{[U]} : T_{[U]} \text{Gr}(n, p) \to \text{Gr}(n, p), \quad \Delta \mapsto \text{Exp}_{[U]}(\Delta).
\]
The associated geodesic is \( t \mapsto [U](t) = \text{Exp}_{[U]}(t\Delta) \), see Fig. 3.1 for an illustration. An explicit formula for Riemannian exponential and thus for the geodesics on the Grassmannian was derived in [8] \( \S 2.5.1 \). For a given pair of initial values \( U \in Gr(n, p) \), \( \Delta \in T_{[U]}Gr(n, p) \), the corresponding geodesic is

\[
t \mapsto \left[ U\Psi \cos(tS)\Psi^T + \Phi \sin(tS)\Psi^T \right] \in Gr(n, p), \quad \Phi S \Psi^T \overset{\text{SVD}}{=} \Delta,
\]

where \( \Phi \in St(n, p) \), \( \Psi \in O_p \) and \( S \in \mathbb{R}^{p \times p} \) diagonal.\(^3\)

For a rectangular, full column-rank matrix \( X \in \mathbb{R}^{n \times p} \), the orthogonal projection onto the column span of \( X \) is

\[
\Pi_X : \mathbb{R}^n \to \text{ran}(X), \quad x \mapsto (X(X^TX)^{-1}X^T)x.
\]

An orthonormal basis (ONB) \( \{u^1, \ldots, u^p\} \subset \mathbb{R}^n \) of \( \text{ran}(X) \) gives rise to a matrix \( U = (u^1, \ldots, u^p) \in St(n, p) \) and the orthogonal projection reduces to \( \Pi_X : x \mapsto UU^T x \).

The principal angles (aka canonical angles) \( \theta_1, \ldots, \theta_p \in [0, \frac{\pi}{2}] \) between two subspaces \([U], \tilde{U}\) \( \in Gr(n, p) \) are defined recursively by

\[
\cos(\theta_k) := \frac{\max_{u \in [U]} \|u\| = 1 \max_{v \in [\tilde{U}]} \|v\| = 1 u^T v}{\|u\| \cdot \|v\|}, \quad \theta_k := \arccos(\sigma_k) \in [0, \frac{\pi}{2}], \quad \sigma_k \text{ is the } k\text{-largest singular value of } U^T \tilde{U} \in \mathbb{R}^{p \times p}, \quad \Theta = (\theta_1, \ldots, \theta_p) \in \mathbb{R}^p.
\]

The principal angles can be computed via \( \theta_k := \arccos(\sigma_k) \in [0, \frac{\pi}{2}] \), where \( \sigma_k \) is the \( k \)-largest singular value of \( U^T \tilde{U} \in \mathbb{R}^{p \times p} \) [10] \( \S 6.4.3 \). The Riemannian subspace distance between \([U], \tilde{U}\) \( \in Gr(n, p) \) is

\[
\text{dist}([U], \tilde{U}) := \|\Theta\|, \quad \Theta = (\theta_1, \ldots, \theta_p) \in \mathbb{R}^p,
\]

see [8] \( \S 2.5.1, \S 4.3 \). Here and throughout, \( \| \cdot \| \) denotes the Euclidean norm.

2. Problem statement and review of the state-of-the-art. Let \( a \in \mathbb{R}^n \), \( b \in \mathbb{R}^p \) and consider the rank-one update

\[
X_{\text{new}} = X + ab^T \in \mathbb{R}^{n \times p}.
\]

Suppose that \( X \) has full column-rank \( p \). Let \( X = U\Sigma V^T \) denote the (thin) SVD of \( X \), where \( U \in St(n, p) \), \( V \in O_p \) and \( \Sigma \) is a regular \( p \)-by-\( p \) diagonal matrix. Let

\[
X_{\text{new}} = X + ab^T = U\Sigma V^T + ab^T = U_{\text{new}} \Sigma_{\text{new}} V_{\text{new}}^T
\]

denote the updated (thin) SVD after the rank-one modification. By writing

\[
X_{\text{new}} = U\Sigma V^T + ab^T = (U + ab^T V \Sigma^{-1}) \Sigma V^T,
\]

we see that \( \text{ran}(U_{\text{new}}) = \text{ran}(X_{\text{new}}) = \text{ran}(U + ab^T V \Sigma^{-1}) \). Hence, the rank-one update on \( X \) acts as a rank-one update on \( U \), which can be considered as an ONB matrix representative \( U \in St(n, p) \) for the subspace \( \text{ran}(X) \).

Objective. The task is to find an orthogonal subspace representative \( \tilde{U}_{\text{new}} \in St(n, p) \) such that \([\tilde{U}_{\text{new}}] = [U_{\text{new}}]\), i.e., such that \( \tilde{U}_{\text{new}} \) spans the same subspace as the updated \( U_{\text{new}} \).

Note that if \( a \in \text{ran}(X) \), then there exists a coefficient vector \( \alpha \in \mathbb{R}^p \) such that \( a = X\alpha \). As a consequence, \( \text{ran}(X + ab^T) = \text{ran}(X(I_p + ab^T)) \subset \text{ran}(X) \), i.e., the rank-one modification does not significantly modify the subspace \( \text{ran}(X) \). Therefore, we restrict our considerations to the case \( a \not\in \text{ran}(X) \).\(^2\)

\(^3\)It is understood that \( \cos \) and \( \sin \) act only on the diagonal elements of \( tS \) in eq. 1.3.

\(^2\)The subspace may be deflated though. This happens for example, if \( -\alpha = b = e_i \).
**Review: rank-one adaptations.** The standard way to approach the above objective is via rank-one SVD updates as considered in [3, 4] and is briefly reviewed below.

The scheme of [3] starts as follows: The rank-one update $X + ab^T = U \Sigma V^T + ab^T$ is written in factorized form as

$$X + ab^T = (U, a) \begin{pmatrix} \Sigma & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} V^T \\ b^T \end{pmatrix}$$

(2.1a)

$$= (U, q) \begin{pmatrix} I_p & U^T a \\ 0 & \|q\| \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I_p & 0 \\ b^T V & 0 \end{pmatrix} \begin{pmatrix} V^T \\ 0 \end{pmatrix}$$

(2.1b)

$$= (U, q) \begin{pmatrix} \Sigma + (U^T a \|\tilde{q}\|) \|q\| \\ 0 \end{pmatrix} \begin{pmatrix} I_p & 0 \\ b^T V & 0 \end{pmatrix} \begin{pmatrix} V^T \\ 0 \end{pmatrix}$$

(2.1c)

$$= (U, q) (\Sigma + (U^T a \|\tilde{q}\|) \|q\|) V^T,$$

(2.1d)

where $\tilde{q} = (I - UU^T)a \neq 0$ is the orthogonal component of $a$ with respect to the subspace $[U]$. 

Note that the left and rightmost matrix factors in the decomposition (2.1d) are (column-)orthogonal by construction. Hence, the updated SVD is obtained by computing the SVD of the (usually small) matrix $K = U\Sigma V^T \in \mathbb{R}^{(p+1) \times p}$ and setting $U_{\text{new}} = (U, q) U'$ as well as $V_{\text{new}} = V V'$. It is worth noting that the procedure works by taking a detour via a representative $(U, q) \in St(n, p+1)$ of a $(p+1)$-dimensional subspace that is pulled back to $St(n, p)$ by the factor $U' \in St(p+1, p)$ from the SVD of $K$. Mind also that an auxiliary numerical computation of the SVD of the $(p+1)$-by-$p$-matrix $K$ is required.

In this work, the focus is on the updated subspace $[U_{\text{new}}] \in Gr(n, p)$. Hence, for obtaining the subspace $[U_{\text{new}}]$, a QR-decomposition of $K$ can be conducted as an alternative to the SVD of $K$. Rank-one adaptations of the QR-decomposition are investigated in [6]. The idea is analogous to the one outlined above: The procedure of [6] p. 775 also starts with a detour via $p+1$ columns: Let $X = QR$ with $Q \in St(n, p)$:

$$X + ab^T = (Q, a) \begin{pmatrix} R \\ b^T \end{pmatrix} = (Q, q) \begin{pmatrix} R \\ b^T \end{pmatrix} + \begin{pmatrix} (Q^T a) \\ \|q\| \end{pmatrix} b^T,$$

(2.2)

where, as before, $\tilde{q} = (I - QQ^T)a$, $q = \frac{\tilde{q}}{\|\tilde{q}\|}$. The algorithm proceeds with applying a suitable sequence of Givens rotations to reestablish the QR-decomposition.

In regards of the work at hand, it is important to emphasize that both the SVD-based approach of [3] and the QR-based approach of [6] make the same detour via $p+1$ columns and that both require an algorithmic decomposition of the auxiliary matrices $K$ and $\tilde{K}$, respectively. To the best of the author’s knowledge, there is no closed-form solution to these subproblems. The original approach of this work eventually avoids the $(p+1)$-columns detour and produces a closed formula for the updated subspace.

### 3. General geometric rank-one subspace adaptation.

In this section, we present a geometric approach to the rank-one subspace adaptation problem introduced in Section 2.

First, we formalize the notion of rank-one subspace adaptations. Just as the notion of a subspace itself, this concept should not depend on the matrix representatives.

DEFINITION 3.1. Let \( \mathcal{U}, \tilde{\mathcal{U}} \in \text{Gr}(n, p) \) be two subspaces of dimension \( p \) in \( \mathbb{R}^n \). We say that \( \mathcal{U} \) and \( \tilde{\mathcal{U}} \) differ by a non-trivial rank-one modification, if all but one of the principle angles between \( \mathcal{U} \) and \( \tilde{\mathcal{U}} \) are zero., i.e., if \( \theta_1 = \cdots = \theta_{p-1} = 0, \theta_p > 0 \).

On the level of matrix subspace representatives, this corresponds indeed to rank-one matrix modifications, as the next lemma shows.

LEMMA 3.2. Two subspaces \( \mathcal{U}, \tilde{\mathcal{U}} \in \text{Gr}(n, p) \) differ by a non-trivial rank-one modification if and only if there exist subspace representatives \( U, \tilde{U} \in S(n, p) \) with \( [U] = \mathcal{U}, [\tilde{U}] = \tilde{\mathcal{U}} \) and non-zero vectors \( x \in \mathbb{R}^n, y \in \mathbb{R}^p \) such that \( \tilde{U} = U + xy^T \).

Proof. ‘\( \Rightarrow \)’ Let \( U, \tilde{U} \in S(n, p) \) be arbitrary matrix representatives of \( \mathcal{U} \) and \( \tilde{\mathcal{U}} \), respectively. Then, \( U^T \tilde{U} \overset{\text{SYD}}{=} RSR^T \), where \( R, \tilde{R} \in O_p \) and \( S = \text{diag}(1, \ldots, 1, s_p), 0 \leq s_p < 1 \), so that \( \theta_p = \arccos(s_p) > 0 \) is the only non-zero principal angle. W.l.o.g., we replace the representatives \( U, \tilde{U} \) with \( UR, \tilde{U} \), so that after this coordinate change, we have

\[
U^T \tilde{U} = (\langle u^j, \tilde{u}^k \rangle)_{j,k=1,\ldots,p} \overset{S}{=} \begin{bmatrix} I_{p-1} \\ s_p \end{bmatrix}.
\]

In particular, by Cauchy-Schwartz’ inequality, \( 1 = \langle u^k, \tilde{u}^k \rangle \leq \| u^k \| \| \tilde{u}^k \| = 1, k = 1, \ldots, p-1 \) and, as a consequence \( u^k = \tilde{u}^k \) for \( k = 1, \ldots, p - 1 \). In summary, \( [U] = \mathcal{U}, [\tilde{U}] = \tilde{\mathcal{U}} \) and

\[
\tilde{U} = U + xy^T, \text{ with } x = (\tilde{u}^p - u^p), \quad y^T = e_p^T = (0, \ldots, 0, 1).
\]

Observe that \( U^T x = (0, \ldots, 0, s_p - 1)^T \neq 0 \).

‘\( \Leftarrow \)’ Suppose that \( x \in \mathbb{R}^n \setminus \{0\}, y \in \mathbb{R}^p \setminus \{0\} \) are such that \( \tilde{U} = U + xy^T \in S(n, p) \). First, note that necessarily \( U^T x \neq 0 \), since otherwise we would have \( I_p = U^T \tilde{U} = I_p + \| x \|^2 yy^T \) and thus \( x = 0 \) or \( y = 0 \). Since both \( U \) and \( \tilde{U} \) have orthonormal columns, it holds \( \|U^T \tilde{U}\| \leq \|U\| \|\tilde{U}\| = 1 \). The principal angles between the subspaces \( [U] \) and \( [\tilde{U}] \) are determined by the singular values of \( U^T \tilde{U} = I_p + U^T xy^T \), which is a non-trivial rank-one modification of the \( (p \times p) \) identity matrix. By [13, Theorem 1], the singular values \( \beta_1 \geq \cdots \geq \beta_p \) of \( B := U^T \tilde{U} \) are sandwiched between the singular values of \( I_p \).
in the following way:

\[ \alpha_k - 1 \geq \beta_k \geq \alpha_{k+1}, \quad k = 1, \ldots, p, \]

where in the case at hand, \( \alpha_0 = \infty, \alpha_1 = \cdots = \alpha_p = 1, \alpha_{p+1} = 0 \). Combined with the fact that \( \beta_1 = \|U^T \hat{U}\| \), this entails \( \beta_k = 1, k = 1, \ldots, p-1 \) and \( 1 \geq \beta_p \geq 0 \). In fact, \( 1 > \beta_p \), for otherwise, \( B = I_p + U^T x y^T \) would be an orthogonal matrix, which can only hold if \( U^T x = 0 \) or \( y = 0 \). As a consequence, the principal subspace angles are

\[ \theta_k = \arccos(\beta_k) = 0, \quad k = 1, \ldots, p-1, \quad \text{and} \quad \frac{\pi}{2} \geq \theta_p = \arccos(\beta_p) > 0. \]

Note that the implicit requirement that \( \hat{U} = U + x y^T \in St(n, p) \), i.e., that the rank-one update by \( x y^T \) preserves the mutual orthonormality of the columns of \( \hat{U} \), imposes special constraints on the selection of the vectors \( x, y \). Intuitively, we expect that any two subspaces \( \text{ran}(X) \) and \( \text{ran}(X + ab^T) \) differ by a rank-one modification in the sense of Definition 3.1. This can be deduced as an immediate consequence of Lemma 3.2 and the upcoming Theorem 3.5. Yet, one can also establish this directly: For example, the QR-update procedure (2.2) gives

\[ X + ab^T = QR + ab^T = (Q, q) \tilde{K} = (Q, q) \left( \begin{array}{c} \phi \\ \varphi^T \end{array} \right) \tilde{R} = (Q\phi + q\varphi^T)\tilde{R}. \]

Here, \( \Phi = \left( \begin{array}{c} \phi \\ \varphi^T \end{array} \right) \in St(p + 1, p) \), so that \( (Q\phi + q\varphi^T) \in St(n, p) \) but \( Q\phi \notin St(n, p) \) in general. The principal angles are determined by the singular values of \( Q^T(Q\phi + q\varphi^T) = \phi \). The singular values of \( \phi \) are the square roots of the eigenvalues of \( \phi^T \phi = I_p - \varphi\varphi^T \). By [3] Theorem 1, all eigenvalues and thus all singular values are equal to 1 with the smallest one as the only exception. Thus, we have proved

**Corollary 3.3.** Two subspaces ran(\( X \)) and ran(\( X + ab^T \)) differ by a rank-one modification in the sense of Definition 3.1. In particular, there exist a subspace representative \( U \) and vectors \( x, y \) such that both \( U \in St(n, p) \) and \( U + x y^T \in St(n, p) \) and, in addition, \( [U] = \text{ran}(X) \), \( [U + x y^T] = \text{ran}(X + ab^T) \).

How to actually obtain such structure-preserving vectors \( x, y \) from a given arbitrary rank-one update in closed form is an alternative way to state the main objective of this work.

We now turn to the geometric solution of the rank-one subspace adaptation problem of Section 2. The idea is to find a geodesic path on the Grassmann manifold \( Gr(n, p) \) that connects \( [U] \) and \( [U_{\text{new}}] \). As outlined in Section 1.4, such a geodesic is determined by a starting point \( [U] \) and a starting velocity \( \Delta \in T_{[U]} Gr(n, p) \) which results in the expression (1.3). In the special case, where the tangent velocity \( \Delta \) is a rank-one matrix \( \Delta = dv^T \), the compact SVD of \( \Delta \) is

\[ \Delta = \Phi \left( \begin{array}{c} s \\ 0_{p-1} \end{array} \right) \Psi^T = \frac{d}{\|d\|} \left( \frac{\|d\|}{\|v\|} \right) v^T =: qsw^T \in \mathbb{R}^{n \times p} \quad (3.1) \]

and the formula for the geodesic (1.3) becomes

\[ t \mapsto \text{Exp}_{[U]}(\Delta) = [U + ((\cos(ts) - 1)Uw + \sin(ts)q) w^T] = [U + \hat{u}(t)w^T]. \quad (3.2) \]

\[^{3}\text{Amongst others, this can be seen from the Sherman-Morrison-Woodbury formula.}\]
This shows that following a geodesic path along a rank-one tangent direction corresponds to a matrix curve of rank-one updates, where $U + \hat{u}(t)w^T \in St(n, p)$ for each $t$. Conversely, this motivates the conjecture that a rank-one adaptation $[U_{new}]$ of a given subspace representative $[U]$ can be reached via a geodesic path along a rank-one tangent direction. The obstacle is to find the associated tangent direction $\Delta$.

This is the main result of this work: Given a rank-one adaptation of a matrix $X + ab^T$ with $\text{ran}(X) = [U]$, $U \in St(n, p)$, we find a rank-one tangent vector $\Delta \in T_U \text{Gr}(n, p)$ of unit norm and a step $t^* \in \mathbb{R}$ such that the associated geodesic crosses the adapted subspace at $t^*$. More precisely,

$$\text{ran}(X + ab^T) = \exp_U(t^* \Delta),$$

with an orthogonal subspace representative $\tilde{U}_{new} \in St(n, p)$, see Fig. 3.1. Note that we can obtain a subspace representative $U \in St(n, p)$ with $\text{ran}(X) = [U]$ from an SVD or a QR-decomposition of $X$. Both alternatives boil down to a decomposition of the form $X = UW$ with $U \in St(n, p)$ and $W \in \mathbb{R}^{p \times p}$ regular and the rank-one update on $X$ leads to a rank-one update on $U$ via $X + ab^T = (U + ab^TW^{-1})W$.

An important building block is the following lemma, which is taken from [14]. It addresses the modification of the orthogonal projector $\Pi_{X_{new}} = U_{new}U_{new}^T$ under the rank-one update on $X$ in closed form. The original lemma addressed the SVD-case of $W = \Sigma V^T$. Adjusted to the general setting of $X = UW$, it reads

**Lemma 3.4 ([14]).** Let $X \in \mathbb{R}^{n \times p}$ feature a decomposition of $X = UW$ with $U \in St(n, p)$ and $W \in \mathbb{R}^{p \times p}$ regular. Let $X_{new} = X + ab^T$ and define

$$\begin{align*}
\tilde{q} &= (I - UU^T)a, \quad q = \frac{\tilde{q}}{\|\tilde{q}\|_2} \in \mathbb{R}^n, \\
g &= \begin{pmatrix} \tilde{w} \\ \omega \end{pmatrix} = \begin{pmatrix} -W^{-T}b \\ \frac{1}{\|\tilde{q}\|_2} (1 + a^TUW^{-T}b) \end{pmatrix} \in \mathbb{R}^{p+1}.
\end{align*}$$

Then the orthogonal projection onto $\text{ran}(X_{new})$ is

$$\Pi_{X_{new}} = (U, q) \begin{pmatrix} U^T \\ q^T \end{pmatrix} - \frac{1}{\|g\|_2^2} (U, q) gg^T \begin{pmatrix} U^T \\ q^T \end{pmatrix}.$$  

(3.4)

For the sake of completeness, a proof of the lemma is included in the appendix.

We are now in a position to state the main theorem.

**Theorem 3.5.** In the same setting as above, consider the rank-one update $X_{new} = X + ab^T = UW + ab^T$. Then,

$$U_{new} = U + (\alpha U w + \beta q) w^T$$

(3.5)

is a valid matrix subspace representative $U_{new} \in St(n, p)$ of the rank-one modified subspace such that $[U_{new}] = \text{ran}(X_{new}) \in \text{Gr}(n, p)$.

---

4If $[U_{new}]$ and $[U]$ are both known, then $\Delta$ can be computed via the Riemannian logarithm, i.e., the inverse of the exponential. The difficulty here is that $[U_{new}]$ is precisely the quantity that is sought after.
The quantities that appear in (3.5) are defined as follows:
\[
\tilde{q} = (I - UU^T)a, \quad q = \frac{\tilde{q}}{\|\tilde{q}\|}, \quad \tilde{w} = -W^{-T}b, \quad w = \frac{\tilde{w}}{\|\tilde{w}\|},
\]
\[
\omega = \frac{1}{\|\tilde{q}\|} (1 - a^T U \tilde{w}) \in \mathbb{R}, \quad g = (\tilde{w}, \omega)^T \in \mathbb{R}^{p+1},
\]
\[
\alpha = \frac{|\omega|}{\|g\|} - 1, \quad \beta = -\text{sign}(\omega) \frac{\|\tilde{w}\|}{\|g\|} \in \mathbb{R}.
\]

In particular, \( \Delta := qw^T \) is a rank-one tangent vector \( \Delta \in T_{[U]} Gr(n, p) \) and the geodesic that starts at \([U] \) with velocity \( \Delta \) meets the point \( \text{ran}(X_{\text{new}}) \) on the Grassmann manifold.

Proof. First, note that by construction \( \Delta^T U = w(q^T U) = 0 \), so that indeed \( \Delta \in T_{[U]} Gr(n, p) \), see (1.2). Consider the corresponding geodesic
\[
t \mapsto [U](t) = \text{Exp}_{[U]}(t\Delta) = [U + ((\cos(ts) - 1)Uw + \sin(ts)q) w^T] \tag{3.7}
\]
From Lemma 3.4, we know that the orthogonal projection onto \( \text{ran}(X_{\text{new}}) \) is
\[
\Pi_{X_{\text{new}}} = UU^T + qq^T - \frac{1}{\|g\|^2} (U\tilde{w} + \omega q)(\tilde{w}^T U + \omega q^T)
\]
\[
= UU^T + \left(1 - \frac{\omega^2}{\|g\|^2}\right) qq^T - \frac{\|\tilde{w}\|^2}{\|g\|^2} (Uww^T U^T) - \frac{\omega\|\tilde{w}\|}{\|g\|^2} (Uwq^T + qw^T U^T).
\]

The geodesic (3.7) leads to a curve of orthogonal projectors \( t \mapsto \Pi_{U(t)} = U(t)U(t)^T \).
An elementary calculation shows that
\[
\Pi_{U(t)} = UU^T + \sin^2(t)qq^T - \sin^2(t) (Uww^T U^T) + \cos(t) \sin(t) (Uwq^T + qw^T U^T).
\]
Comparing the expressions of \( \Pi_{X_{\text{new}}} \) and \( \Pi_{U(t)} \) term by term, we see that the task is reduced to find \( t^* \) such that
\[
\left(1 - \frac{\omega^2}{\|g\|^2}\right) = \sin^2(t^*) = \frac{\|\tilde{w}\|^2}{\|g\|^2}, \quad \text{and} \quad -\frac{\omega\|\tilde{w}\|}{\|g\|^2} = \cos(t^*) \sin(t^*).
\]
This is indeed possible: First, recall that \( g = (\tilde{w}^T, \omega)^T \) and observe that
\[
\left(1 - \frac{\omega^2}{\|g\|^2}\right) = \frac{\|g\|^2 - \omega^2}{\|g\|^2} = \frac{\|\tilde{w}\|^2}{\|g\|^2} < 1.
\]
Hence, \( \sin(t^*) = \pm \frac{\|\tilde{w}\|}{\|g\|} \).
As a consequence,
\[
\cos(t^*) \sin(t^*) = \sqrt{1 - \sin^2(t^*)} \sin(t^*) = \sqrt{\frac{\|g\|^2 - \|\tilde{w}\|^2}{\|g\|^2}} (\pm 1) \frac{\|\tilde{w}\|}{\|g\|}
\]
\[
= \pm \frac{\|\omega\| \|\tilde{w}\|}{\|g\|} \frac{1}{\|g\|} = \frac{\omega \|\tilde{w}\|}{8}.
\]
Hence, the sign must be chosen such that $\sin(t^*) = -\text{sign}(\omega) \frac{\|\tilde{w}\|}{\|g\|}$ and we obtain

$$\alpha = \cos(t^*) - 1 = \frac{\|\omega\|}{\|g\|} - 1, \quad \beta = \sin(t^*) = -\text{sign}(\omega) \frac{\|\tilde{w}\|}{\|g\|}.$$ 

The corresponding step is $t^* = \arcsin \left( -\text{sign}(\omega) \frac{\|\tilde{w}\|}{\|g\|} \right)$. For this choice of $t^*$, the orthogonal projectors $\Pi_{X_{\text{new}}} = \Pi_{\text{U}(t^*)}$ coincide. Due to the one-to-one correspondence between a subspace and the orthogonal projection onto it, we obtain

$$[U](t^*) = \text{ran}(X_{\text{new}}),$$

as claimed. □

**Subspace distance.** The theorem allows us to compute the Riemannian subspace distance between the original subspace $\text{ran}(X) = [U]$ and the adapted subspace $\text{ran}(X + ab^T) = [U_{\text{new}}]$.

**Corollary 3.6.** The Riemannian subspace distance between the original subspace $\text{ran}(X) = [U]$ and the adapted subspace $\text{ran}(X + ab^T) = [U_{\text{new}}]$ is

$$\text{dist}([U], [U_{\text{new}}]) = \arccos \left( \frac{\|\omega\|}{\|g\|} \right) = \arccos \left( \frac{\|\omega\|}{\sqrt{\|\tilde{w}\| + \omega^2}} \right),$$

where $\tilde{w}, \omega$ are as introduced in Theorem 3.5.

**Proof.** Let $w$ be as in the theorem and let $W^\perp$ be an orthogonal completion such that $Z = (W^\perp, w) \in O_p$. From (3.5), we have

$$U^T U_{\text{new}} = U^T (U + (\alpha U w + \beta q) w^T) = I + \alpha w w^T$$

$$= I + \alpha Z \begin{bmatrix} 0_{p-1} \\ 1 \end{bmatrix} Z^T = Z \left( \begin{bmatrix} I_{p-1} \\ 0 \end{bmatrix} + \begin{bmatrix} 0_{p-1} \\ \cos(t^*) - 1 \end{bmatrix} \right) Z^T.$$ 

Hence, the subspace distance is

$$\|\langle \text{arccos}(1), \ldots, \text{arccos}(1), \text{arccos}(\cos(t^*)) \rangle \| = |t^*| = \text{arccos} \left( \frac{\|\omega\|}{\|g\|} \right).$$

(This can also be seen by converting (3.5) back to the general form (1.3)). □

**Recovering $X_{\text{new}}$.** If $X \in \mathbb{R}^{n \times p}$ has a decomposition $X = UW$ with $U \in \text{St}(n, p)$, then Theorem 3.5 gives $U_{\text{new}} \in \text{St}(n, p)$ such that $\text{ran}(X_{\text{new}}) = \text{ran}(X + ab^T) = [U_{\text{new}}]$. We may use this to construct a decomposition $X_{\text{new}} = U_{\text{new}} W_{\text{new}}$. In fact, since $\text{ran}(X + ab^T) = [U_{\text{new}}]$, it holds

$$X_{\text{new}} = X + ab^T = UW + ab^T = U_{\text{new}} W_{\text{new}},$$

with a suitable $W_{\text{new}} \in \mathbb{R}^{p \times p}$. Multiplying with $U_{\text{new}}^T$ from the left gives

$$W_{\text{new}} = U_{\text{new}}^T UW + U_{\text{new}}^T ab^T.$$ 

This is also clear from the fact that $U_{\text{new}}^T X_{\text{new}} = X_{\text{new}}$. By inserting the explicit formula for $U_{\text{new}} = U + (\alpha U w + \beta q) w^T$, the updated $W_{\text{new}}$ is obtained from a rank-one update on $W$ via

$$W_{\text{new}} = W + (U^T a + \gamma w) b^T,$$

$$\gamma = \left( \beta (q^T a) - \alpha \frac{\|q\| \|w\|}{\|\tilde{w}\|} \right) \in \mathbb{R},$$

(3.11)
where all quantities are as introduced in Theorem 3.5. As a consequence, the rank-
one update $X_{new} = UW + ab^T = U_{new}W_{new}$ splits into a orthogonality-preserving
rank-one update on $U$, which gives $U_{new}$ and an associated rank-one update on $W$, which makes the geometric update of the orthogonal decomposition $X = UW$ very
efficient. When compared to the classical SVD- or QR updates, one loses the special
structure of the $W$-factor, though.

**Computational complexity.** Computationally, Theorem 3.5 reduces the rank-
one modified orthogonal decomposition to an elementary matrix update

$$U_{new} = U + xy^T \in \mathbb{R}^{n \times p}.$$  

Computing $x = U(\alpha w) + \beta q$ requires the following operations:

- $\bar{a} = U^T a : \ np \ \text{FLOPS}$,
- $\bar{q} = a - U \bar{a} : \ np + n \ \text{FLOPS}$,
- $q = \frac{\bar{q}}{||\bar{q}||} : \ 2n \ \text{FLOPS}$,
- $\bar{w} = -W^{-T} b, \ w = \frac{\bar{w}}{||\bar{w}||} : \ \mathcal{O}(p^3) \ \text{FLOPS}$,
- $\omega = \frac{1}{||\bar{q}||}(1 - \bar{a}^T \bar{w}) : \ \mathcal{O}(p) \ \text{FLOPS}$,
- $g = (\bar{w}^T, \omega)^T, ||g|| : \ \mathcal{O}(p) \ \text{FLOPS}$,
- $\alpha = \frac{||\omega||}{||g||} - 1 : \ \mathcal{O}(1) \ \text{FLOPS}, \ \beta = -\text{sign}(\omega) \frac{||\bar{w}||}{||g||} : \ \mathcal{O}(1) \ \text{FLOPS}$.

Assuming that $n \gg p$, we count only the operations that scale in $n$: These sum up to
$3np + 4n$ (computing $x$) + $np$ (computing $U + xy^T$) = $4np + 4n = \mathcal{O}(np)$ \ FLOPS.

Note that all the terms that appear in the $W$-update (3.11) are already available
from the computations for $U_{new}$. Therefore, after $U_{new}$ is known, the corresponding
$W_{new}$ is obtained via an elementary rank-one update on the $(p \times p)$-matrix $W$ which
consumes $\mathcal{O}(p^2) \subset \mathcal{O}(np)$ \ FLOPS, see (3.11). Hence, both factors $U_{new}, W_{new}$ of the
complete update of the orthogonal decomposition

$$X_{new} = UW + ab^T = U_{new}W_{new}.$$  

are obtained in $\mathcal{O}(np)$ \ FLOPS\(^5\)

In contrast, the rank-one adaptations (2.1d) and (2.2) require the matrix-matrix
product of an $n \times (p + 1)$ matrix with a $(p + 1) \times p$-matrix for computing the adapted
subspace representative, which alone takes $n(p + 1)p = \mathcal{O}(np^2)$ \ FLOPS.

**Related work.** The so-called Grassmannian Rank-One Update Subspace Estimation (GROUSE, [2]) considers the subspace-dependent, unique residual associated
with a least-squares problem of the form

$$\arg \min_{\alpha \in \mathbb{R}^p} ||A^T U \alpha - a||^2_2$$

as a differentiable function on the Grassmannian

$$F : Gr(n, d) \to \mathbb{R}, \ \ [U] \to a^T (I - A^T U (U^T A A^T U)^{-1} U^T A) a.$$  

GROUSE is an iterative optimization scheme that operates on a sequence of incoming, possible incomplete data vectors $a \in \mathbb{R}^n \setminus \{0\}$. Each iteration is based on following the Grassmann geodesic in the direction of steepest descent $-\nabla_{[U]} F$ with

\(^5\)The Level-2 BLAS operation “DGER” performs the rank-one operation $A := \alpha xy^T + A$.  

10
respect to the above residual norm function $F$. The Grassmann gradient is $\nabla_{[U]} F = -2A(a - ATU\alpha)\alpha^T = -2Ar(\alpha)\alpha^T \in T_{[U]}Gr(n,p)$, where $\alpha = (UTA^TU)^{-1}UTAa$ is the optimal coefficient vector associated with the above least-squares problem and $a - ATU\alpha = r(\alpha)$ is the corresponding residual vector. Note that the gradient is of rank-one.

GROUSE thus makes extensive use of the formulas (3.1), (3.2) for the specific rank-one descent directions $\Delta = qsw^T = 2Ar(\alpha)\alpha^T$. In the simplest case, where $A = I_n$, it holds $q = \frac{(I - UU^T)a}{\|UU^T\|_F}$, $w = \frac{\alpha}{\|\alpha\|} = \frac{Ua}{\|Ua\|}$, which shows that the left-singular vector $q$ and the right-singular vector $w$ are not independent quantities, since both vectors are functions of $a$ and $U$. In any case, the singular vectors $q$ and $w$ from the SVD of the rank-one gradient $\nabla_{[U]} F$ are special in the sense that they correspond to a certain subset of rank-one tangent vectors that arise from least-squares problems, where both exhibit a functional dependency on $a, U$ (and $A$). In contrast, Theorem 3.5 considers completely general rank-one tangent vectors.

**Appendix A. Proof of Lemma**.

*Proof.* We start with a decomposition inspired by [3, eq. (3)]. Note that $(U,q) \in St(n,p + 1)$ by construction. It holds that

$$X + ab^T = (U,q) \left( \begin{array}{c} W + U^T ab^T \\ \|q\|b^T \end{array} \right) =: (U,q)K,$$

where $K \in \mathbb{R}^{(p+1)\times p}$. Let $K = \hat{U}\hat{W}$ be a decomposition of $K$ with $\hat{U} \in St(p + 1, p)$, $\hat{W} \in \mathbb{R}^{p\times p}$ regular. Hence,

$$X + ab^T = \left( (U,q)\hat{U} \right) \hat{W} =: U_{new}W_{new}.$$

Let $g \in \mathbb{R}^{p+1}$ be such that $(\hat{U}, \frac{g}{\|g\|}) \in O_{p+1}$ is an orthogonal completion of $\hat{U}$. Because of $I_{p+1} = (\hat{U}, \frac{g}{\|g\|})(\hat{U}, \frac{g}{\|g\|})^T$, we have

$$\hat{U}\hat{U}^T = I_{p+1} - \frac{1}{\|g\|^2}gg^T$$

and, as a consequence,

$$U_{new}U_{new}^T = (U,q)\hat{U}\hat{U}^T \left( \begin{array}{c} UT^T \\ q^T \end{array} \right) = (U,q) \left( I_{p+1} - \frac{1}{\|g\|^2}gg^T \right) \left( \begin{array}{c} UT^T \\ q^T \end{array} \right).$$

Hence, it is sufficient to determine $g$, which is characterized up to a scalar factor by $\hat{U}^Tg = 0$. Since $\text{ran}(K) = \text{ran}(\hat{U}^Tg = 0)$, this condition is equivalent to $K^Tg = 0$. Let $\hat{w} \in \mathbb{R}^p$ denote the first $p$ components of $g$ and let $\omega \in \mathbb{R}$ be the last entry such that $g^T = (\hat{w}^T,\omega)$. When writing the equation $g^TK = 0$ as

$$(\hat{w}^T,\omega) \left( \begin{array}{cc} I_p & UTa \\ 0 & \|q\|_2 \end{array} \right) \left( \begin{array}{c} W \\ b^T \end{array} \right) = 0,$$

it is straightforward to show that $g = \left( \frac{1}{\|q\|_2}(-W^{-T}b) \right) \in \mathbb{R}^{p+1}$ and any scalar multiple of this vector is a valid solution. Using this vector in (A.1) proves the lemma.

**Appendix B. MATLAB code.**
function [ Unew, Wnew, Sdist ] = grood(U, W, a, b)
% MATLAB function for performing the adaptation of Theorem 3.5.
%-------------------------------------------------------------
% Perform a (G)eometric (R)ank-(O)ne (O)rthogonal (D)ecomposition
% Unew Wnew = UW + ab'
% on a matrix X with known orthogonal decomposition X=UW
%
% Inputs:
% UW: decomposition of X=UW into a column-orthogonal
% (nxp)-matrix U and a regular (pxp)-matrix W
% a: n-vector
% b: p-vector
%
% Outputs:
% Unew: column-orthogonal matrix with range
% ran(Unew) = ran(U+ab')
% Wnew: second factor in orthogonal matrix decomposition
% Sdist: subspace distance between ran(Unew) and ran(U)
%
% Remark: if a is in ran(U), the algorithm returns U, W
%
% author: R: Zimmermann, IMADA, SDU Odense
% zimmermann@imada.sdu.dk
%-------------------------------------------------------------
% compute orthogonal component of a and normalize
Ua= U'*a;
q = a - U*Ua;
n_q = norm(q);
q_n = q/n_q;
%
% proceed only, if orth. component is significantly nonzero
if n_q > 1.0e-10
% modify b
w = linsolve(W', -b);
% Grassmann update:
omega = (1.0/n_q)*(1 - Ua'*w);
n_w = norm(w);
w_n = w/n_w;
% compute the norm of g = (w, omega)
q = sqrt(n_w^2 + omega^2);
% compute the (cos(t)-1) and sin(t)-factors:
sin_factor = -sign(omega)*(n_w/n_q);
cos_factor = abs(omega)/n_w - 1;
% compute the rank one update
rank1_up = (cos_factor*U*w_n + sin_factor*q_n)*w_n';
Unew = U + rank1_up;
% compute the update on W
Wnew = W + (Ua +... (sin_factor*(q_n*a) - cos_factor*(n_q*omega)/n_w)*w_n)*b';
%subspace distance
Sdist = acos(abs(omega)/n_g);
else
    Unew = U;
    Wnew = W;
    Sdist = 0.0;
return;
end

**MATLAB script performing a random rank-one update.**

```
% Numerical experiment
% @author: Ralf Zimmermann, IMADA, SDU Odense
% zimmermann@imada.sdu.dk
%
clear;
format long
%
% set dimensions
n = 15000;
p = 500;
% create random data
X = rand(n,p);
a = rand(n,1);
b = rand(p,1);

% initial QR
[Q,W] = qr(X, 0);
% for comparison purposes: brute-force update
Xbrute = X + a*b';
[Qbrute, Wbrute] = qr(Xbrute, 0);
% check distance between Q and Qnew
dist_Q_Qbrute = subspaceDist(Q, Qbrute);

% perform geometric rank-one orthogonal decomposition
% compute subspace distance according to Cor. 3.5 on-the-fly
[Qnew, Wnew, subspace_dist_cor35] = grood(Q, W, a, b);

disp(['Distance between the original and the adapted subspace:'])
disp(['- computed via brute force : ', num2str(dist_Q_Qbrute, 14), '.'])
disp(['- computed via Corollary 3.5: ', num2str(subspace_dist_cor35, 14), '.'])

% check if Qnew indeed spans the same subspace
% as the brute-force Qbrute
dist_Qnew_Qbrute = subspaceDist(Qnew, Qbrute);
disp(['The subspaces computed by brute force and by Theorem 3.4'])
disp(['match up to a subspace distance of: ', num2str(dist_Qnew_Qbrute, 14)])
```
% reconstruct X + ab’ = QW + ab’ = Qnew Wnew:
Xnew = Qnew*Wnew;
norm_Xbrute_Xnew = norm(Xbrute - Xnew);
disp(['The actual Xnew = X + ab and the reconstruction Qnew Wnew'])
disp(['match up to a numerical accuracy of: ', ...
     num2str(norm_Xbrute_Xnew, 14)])

Auxiliary function for computing the distance between two subspaces.

function [ dist ] = subspaceDist(U0, U1)
%-------------------------------------------------------------
% compute Grassmann distance dist(U0,U1)
% of to subspaces represented by U0 and U1
% in terms of the norm of the canonical angles
% see, e.g., Golub/Van Loan
%
% author: R: Zimmermann, IMADA, SDU Odense
% zimmermann@imada.sdu.dk
%-------------------------------------------------------------
[Q, S, R] = svd((U0'*U1));

% catch numerical round off errors:
% enforce that all singular vals are <=1
s = diag(S);
for k = 1:length(s)
    if abs(s(k) - 1.0) < 100*eps || s(k)>1.0
        s(k) =1.0;
    end
end
theta = acos(s);
dist = norm(theta,2);
return;
end

REFERENCES
[1] P.-A. Absil, R. Mahony, and R. Sepulchre. Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, New Jersey, 2008.
[2] L. Balzano, R. Nowak, and B. Recht. Online identification and tracking of subspaces from highly incomplete information. In Communication, Control, and Computing (Allerton), 48th Annual Allerton Conference on, pages 704–711. IEEE, 2010.
[3] M. Brand. Fast low-rank modifications of the thin singular value decomposition. Linear Algebra and its Applications, 415:20–30, 2006.
[4] J. R. Bunch and C. P. Nielsen. Updating the singular value decomposition. Numerische Mathematik, 31:111–129, 1978.
[5] J. R. Bunch, C. P. Nielsen, and D. C. Sorensen. Rank-one modifications of the symmetric eigenproblem. Numerische Mathematik, 31:31–48, 1978.
[6] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. Mathematics of Computation, 30:772–795, 1976.
[7] M.P. do Carmo and F. Flaherty. Riemannian Geometry. Mathematics: Theory & Applications. Birkhäuser Boston, 2013.
[8] A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. SIAM Journal on Matrix Analysis and Application, 20(2):303–353, 1999.
[9] P. E. Gill, G. H. Golub, W. Murray, and M. A. Saunders. Methods for modifying matrix factorizations. *Mathematics of Computation*, 28:505–535, 1974.

[10] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, 4 edition, 2013.

[11] M. Gu and S. C. Eisenstat. A stable and efficient algorithm for the rank-one modification of the symmetric eigenproblem. *SIAM Journal on Matrix Analysis and Applications*, 15:1266–1276, 1994.

[12] M. Moonen, P. Van Dooren, and J. Vandewalle. A singular value decomposition updating algorithm for subspace tracking. *SIAM Journal on Matrix Analysis and Applications*, 13(4):1015–1038, 1992.

[13] R. C. Thompson. The behavior of eigenvalues and singular values under perturbations of restricted rank. *Linear Algebra and its Applications*, 13:69–78, 1976.

[14] R. Zimmermann, B. Peherstorfer, and K. Willcox. Geometric subspace updates with applications to online adaptive nonlinear model reduction. *SIAM Journal on Matrix Analysis and Application*, 39(1):234–261, 2018.