Procrustes problems in finite dimensional indefinite scalar product spaces

Ulric Kintzel

Institut für Mathematik, MA 4-5, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

Received 27 May 2003; accepted 18 January 2005
Available online 2 March 2005
Submitted by V. Mehrmann

Abstract

Some applications from multidimensional scaling in an environment of indefinite scalar products are investigated. In particular, the construction of points from given distances is considered, and two variants of Procrustes problems are discussed. It is found that both Procrustes problems can be solved with the help of $H$-polar decompositions or $(G,H)$-polar decompositions of specified matrices. Furthermore, a method for the numerical computation of $H$- or $(G,H)$-polar decompositions of nonsingular matrices is described.

© 2005 Elsevier Inc. All rights reserved.

AMS classification: 15A63; 15A23

Keywords: Indefinite scalar products; Polar decompositions; Procrustes problems

1. Introduction

Let $\mathbb{F}$ be the field of real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$ and let $\mathbb{F}^n$ be an $n$-dimensional vector space over $\mathbb{F}$. Furthermore, let $H$ be a fixed chosen nonsingular symmetric or Hermitian matrix in $\mathbb{F}^{n \times n}$ and let $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_n)^T$ be column vectors in $\mathbb{F}^n$. Then the bilinear or sesquilinear functional
\[ [x, y] = (Hx, y) \] where \((x, y) = \sum_{\alpha=1}^{n} x^{\alpha} y^{\alpha} (y^{\alpha} = y^{\alpha} \text{ if } F = \mathbb{R})\)
defines an indefinite scalar product in \(\mathbb{F}^n\). Indefinite scalar products have almost all the properties of ordinary scalar products, except for the fact that the value of \([x, x]\) for a vector \(x \neq 0\) can be positive, negative or zero. A corresponding vector is called positive (space-like), negative (time-like) or neutral (isotropic, light-like), respectively. If \(A\) is an arbitrary matrix in \(\mathbb{F}^{n \times n}\), then its \(H\)-adjoint \(A^*[\cdot]\) is characterised by the property that
\[ [Ax, y] = [x, A^*[\cdot]]y \] for all \(x, y \in \mathbb{F}^n\).
This is equivalent to the fact that between the \(H\)-adjoint \(A^*[\cdot]\) and the ordinary adjoint \(A^* = A^T\) there exists the relationship
\[ A^*[\cdot] = H^{-1} A^* H. \]
If in particular \(A^*[\cdot] = A\) or \(A^* H = HA\), one speaks of an \(H\)-selfadjoint or \(H\)-symmetric or \(H\)-Hermitian matrix, and an invertible matrix \(U\) with \(U^*[\cdot] = U^{-1}\) or \(U^* H U = H\) is called an \(H\)-isometry or an \(H\)-orthogonal or \(H\)-unitary matrix. If \(A\) is a given matrix in \(\mathbb{F}^{n \times n}\), then a factorisation of the form
\[ A = UM \] with \(U^* H U = H\) and \(M^* H = HM\)
is called an \(H\)-polar decomposition of \(A\).

\(H\)-polar decompositions are an important tool for this paper, in which two applications from a branch of mathematics, known in psychology as factor analysis or multidimensional scaling (MDS), are generalised into the environment of indefinite scalar products. In a typical application of MDS (for example see [1]) test persons are first requested to estimate the dissimilarity (or similarity) of specified objects which are selected terms describing the subject of the analysis. In this way the comparison of \(N\) objects in pairs produces the similarity measures, called proximities, \(p_{kl}, 1 \leq k, l \leq N\), from which the distances \(d_{kl} = f(p_{kl})\) are then determined using a function \(f\), for example \(f(x) = ax + b\), which is called the MDS model. Using these distances, the coordinates of points \(x_k\) in an \(n\)-dimensional Euclidean space are constructed such that \(\|x_k - x_l\| = d_{kl}\) where \(\|\cdot\|\) stands for the Euclidean norm. Thus each object is represented by a point in a coordinate system and the data can be analysed with regard to their geometric properties.

The results of interrogating the test persons are often categorised in groups, producing several descriptive constellations of points which must be mutually compared in the analysis. To make such a comparison of two constellations \(x_k\) and \(y_k\) possible, it is first of all necessary to compensate for irrelevant differences resulting from possibly different locations in space. This is done with an orthogonal transformation \(U\) devised such that \(\sum_k \|Ux_k - y_k\|^2\) is minimised. Thereafter the constellations \(x'_k = Ux_k\) and \(y'_k\) can be analysed.

The MDS model \(f\) is chosen in particular by adding a constant \(b\) (and by making further assumptions such as \(d_{kk} = 0\)), so that the triangle inequality is fulfilled and
therefore the points can be embedded in a Euclidean space [1, Chapter 18]. However, this means that the transformed data $d_{kl}$ describe completely different geometric properties than the original data $p_{kl}$ do. This is the starting point of this paper in which the stated tasks are considered in the following way:

Section 3 describes a method for constructing vectors $x_k$ and an indefinite scalar product $.. = (H, ..)$ such that $x_k - x_l, x_k - x_l = q_{kl}$ where $q_{kl}$ are given real numbers, for example $q_{kl} = p_{kl}^2$. In Section 4 the problem of finding an $H$-isometry $U$ for which $x_k' = Ux_k$ and $y_k$ are at optimum congruence will be discussed. This so-called Procrustes problem$^1$ does not always admit a solution, so that an alternative approach in Section 5 will be considered. The investigation of the Procrustes problems requires some background on $H$-polar decompositions which will be provided in the preparatory Section 2. Furthermore, the final Section 6 describes a method for the numerical computation of $H$-polar decompositions of nonsingular matrices which intends to make the practical application of the results on the Procrustes problems possible, too.

The following notation is used: The rank of a matrix $A$ is denoted by rank $A$. If the matrix $A$ is square, then tr $A$, det $A$ and $\sigma(A)$ are its trace, determinant and spectrum, respectively. Furthermore, the abbreviation $A^{-*} = (A^*)^{-1} = (A^{-1})^*$ is used. $I_p$, $Z_p$ and $J_p(\lambda)$ denote the $p \times p$ identity matrix, the $p \times p$ matrix with ones on the anti-diagonal and otherwise zeros, and the $p \times p$ upper Jordan block for the eigenvalue $\lambda$, respectively. Moreover, $A_1 \oplus \cdots \oplus A_k$ represents the block diagonal matrix consisting of the specified blocks, and $\text{diag}(\alpha_1, \ldots, \alpha_k)$ stands for a diagonal matrix with the specified diagonal elements. Even when no further specifications are made, a nonsingular (real) symmetric or (complex) Hermitian matrix is always meant by $H$, and instead of $A[\ast]$ we sometimes write $A^H$ to indicate the underlying scalar product.

2. Polar decompositions in indefinite scalar products

$H$-polar decompositions of real or complex matrices have been investigated in detail in [2–4,14] as well as in the further references specified there. An essential result of these investigations is the following fact.

**Proposition 2.1** ($F = \mathbb{R}$ or $F = \mathbb{C}$). A matrix $A \in \mathbb{F}^{n \times n}$ admits an $H$-polar decomposition if and only if there exists an $H$-selfadjoint matrix $M$ such that $M^2 = A[\ast]A$ and $\ker M = \ker A$. Moreover, such a matrix $M$ always exists if $\sigma(A[\ast]A) \subset \mathbb{C} \setminus (-\infty, 0]$ and can then be chosen such that $\sigma(M) \subset \{z \in \mathbb{C} | \text{Re}(z) > 0 \}$.

---

$^1$ Procrustes, a robber in Greek mythology, who lived near Eleusis in Attica. Originally he was called Damastes or Polypemon. He was given the name Procrustes (“the stretcher”) because he tortured his victims to fit them into a bed. If they were too tall, he chopped off their limbs or formed them with a hammer. If they were too small, he stretched them. He was overcome by Theseus who served him the same fate by chopping off his head to fit him into the bed.
Proof. See [2, Theorem 4.1, Theorem 4.4 and Lemma 7.8]. □

If the matrix $A^{[*]}A$ has non-positive real eigenvalues, the matrix $M$ from the proposition can only exist, when the canonical form [6, Theorem I.3.3]

$$S^{-1}A^{[*]}AS = \bigoplus_{i=1}^{r} J_{p_i}(\lambda_i) \oplus \bigoplus_{i=r+1}^{s} \begin{bmatrix} J_{p_i}(\lambda_i) & \vdots \\ \vdots & J_{p_i}(\lambda_i) \end{bmatrix},$$

$$S^{[*]}HS = \bigoplus_{i=1}^{r} \varepsilon_i Z_{p_i} \oplus \bigoplus_{i=r+1}^{s} \begin{bmatrix} Z_{p_i} & Z_{p_i} \end{bmatrix},$$

$$\lambda_1, \ldots, \lambda_r \in \mathbb{R}, \varepsilon_1, \ldots, \varepsilon_r \in \{+1, -1\}$$

and

$$\lambda_{r+1}, \ldots, \lambda_s \in \mathbb{C} \setminus \mathbb{R}$$

of the pair $(A^{[*]}A, H)$ satisfies certain conditions ([2, Theorem 4.4] and [4, Errata]). One of these conditions says, that the part of the canonical form belonging to a negative eigenvalue $\lambda \in \sigma(A^{[*]}A)$ must be of the form

$$(k \bigoplus_{i=1}^{r} J_{p_i}(\lambda) \oplus J_{p_i}(\lambda)), k \bigoplus_{i=1}^{r} [Z_{p_i} \oplus -Z_{p_i}],$$

(2.1)

Further conditions must be satisfied if $A^{[*]}A$ is singular.

A matrix $M$ is said to be $H$-nonnegative, if $HM$ is positive semidefinite. The particular $H$-polar decompositions in which the matrix $M$ is $H$-nonnegative are called semidefinite $H$-polar decompositions [3, Section 5]. The criteria for the existence of such a decomposition are described with the next result.

**Proposition 2.2** ($\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$). A matrix $A \in \mathbb{F}^{n \times n}$ admits a semidefinite $H$-polar decomposition if and only if the following conditions are satisfied:

(i) The canonical form of the pair $(A^{[*]}A, H)$ is of the kind

$$S^{-1}A^{[*]}AS = \bigoplus_{j=1}^{k} \begin{bmatrix} \lambda_j I_{p_j} & \vdots \\ \vdots & \lambda_j I_{q_j} \end{bmatrix} \oplus \begin{bmatrix} 0_{r+s} & 0_{r+t} \\ 0_{r+t} & 0_{r+s} \end{bmatrix},$$

$$S^{[*]}HS = \bigoplus_{j=1}^{k} \begin{bmatrix} I_{p_j} & \vdots \\ \vdots & -I_{q_j} \end{bmatrix} \oplus \begin{bmatrix} I_{r+s} & 0_{r+t} \\ 0_{r+t} & -I_{r+s} \end{bmatrix},$$

where $\lambda_j > 0$, $p_j, q_j \in \mathbb{N}$ for $1 \leq j \leq k$ and $r, s, t \in \mathbb{N}$.

(ii) There exists a basis \{s_1, \ldots, s_{n-2r-s-t}, e_1, \ldots, e_{r+s}, f_1, \ldots, f_{r+t}\} in which these blocks appear and for which

$$\ker A = \operatorname{span}(e_1 + f_1, \ldots, e_r + f_r) \oplus \operatorname{span}(e_{r+1}, \ldots, e_{r+s}) \oplus \operatorname{span}(f_{r+1}, \ldots, f_{r+t}).$$

Proof. See [3, Theorem 5.3]. □
This paper will also make use of indefinite polar decompositions where the factors $U$ and $M$ are doubly structured with respect to two selfadjoint matrices $G$ and $H$.

**Definition 2.3** ($F = \mathbb{R}$ or $F = \mathbb{C}$). Let $G, H \in \mathbb{F}^{n \times n}$ be nonsingular and selfadjoint and let $A \in \mathbb{F}^{n \times n}$. A factorisation of the form $A = UM$ with $UH = UG = U^{-1}$ and $MH = MG = M$ is called a $(G, H)$-polar decomposition of $A$. A matrix having the properties of $U$ is said to be $(G, H)$-isometric (-orthogonal or -unitary), and a matrix having the properties of $M$ is said to be $(G, H)$-selfadjoint (-symmetric or -Hermitian). If the factor $M$ in particular is $H$-nonnegative ($HM \geq 0$), the factorisation is called an $H$-semidefinite $(G, H)$-polar decomposition.

These factorisations will be of interest in the special case in which the matrices $G$ and $H$ satisfy

$$H^{-1}G = \mu^2 G^{-1}H$$

for some $\mu \in \mathbb{R} \setminus \{0\}$. (2.2)

A pair of matrices which has this property can be characterised as follows.

**Lemma 2.4** ($F = \mathbb{R}$ or $F = \mathbb{C}$). Let $G, H \in \mathbb{F}^{n \times n}$ be nonsingular and selfadjoint. Then (2.2) is satisfied if and only if there exists a nonsingular matrix $S \in \mathbb{F}^{n \times n}$ such that $S^*HS = I_p \oplus -I_q \oplus I_r \oplus -I_s$ and $S^*GS = \mu (I_p \oplus -I_q \oplus -I_r \oplus I_s)$ for suitable constants $p, q, r, s \in \mathbb{N}$ with $p + q + r + s = n$.

**Proof.** $[\Rightarrow]$: Let $A = \mathbb{F}^{n \times n}$ be a nonsingular matrix such that $A = \mu^2 A^{-1}$ for some $\mu \in \mathbb{R} \setminus \{0\}$. Then $A^2 = \mu^2 I$ so that the Jordan normal form of $A$ must have the form $P^{-1}AP = J = \text{diag}(\pm \mu)$. In particular, if $A = H^{-1}G$, it follows that $(P^*HP)^{-1}(P^*GP) = P^{-1}H^{-1}GP = J^* = (P^*GP)(P^*HP)^{-1}$. Thus the selfadjoint matrices $P^*HP$ and $P^*GP$ commute and can therefore be diagonalised simultaneously, so that an orthogonal or unitary matrix $Q$ consisting of eigenvectors of $P^*HP$ (or $P^*GP$) can now be chosen for which $P^*HP = QA_HQ^*$ and $P^*GP = QA_GQ^*$, where $A_H$, $A_G$ are diagonal matrices containing the real eigenvalues. This means that $(A_H^{-1}A_G)^2 = (Q^*(QA_HQ^*)^{-1}(QA_GQ^*)Q)^2 = (Q^*JQ)^2 = \mu^2 I$ and consequently $A_H^{-1}A_G$ can also be written in the form $^2$

$^2$ The matrices $\mu \Sigma$ and $J$ have the same diagonal elements, but their ordering may be different.
Setting $A_H^{-1}A_G = \mu \Sigma$ with $\Sigma = \text{diag}(\pm 1)$.

Hence, for $S = S$ such that $(J_H)$-semidefinite if and only if both $\Sigma_H = \text{sign}(A_H)$, $\Sigma_G = \text{sign}(A_G)$, $\epsilon = \text{sign}(\mu)$, we obtain

$$|A_H|^{-1}|A_G| = |\mu|I \quad \text{and} \quad \Sigma_H \Sigma_G = \epsilon \Sigma.$$

The asserted form can always be obtained by suitable permutation. (The operations on $A$ are to be applied to its diagonal elements.)

$|\Leftarrow|$: The assertion follows directly from $H^{-1}G = \mu S(I_{p+q} \oplus -I_{r+s})S^{-1}$ and $G^{-1}H = \mu^{-1}S(I_{p+q} \oplus -I_{r+s})S^{-1}$. \hfill $\square$

Obviously, a $(G, H)$-polar decomposition of a matrix $A$ can exist only if

$$H^{-1}A^*H = H^{-1}M^*HH^{-1}U^*H = G^{-1}M^*GG^{-1}U^*G = G^{-1}A^*G$$

or $A^H = A^G$. These matrices allow the following representation.

Lemma 2.5 (F $\in \mathbb{R}$ or $F = \mathbb{C}$). Let $G, H \in \mathbb{F}^{n \times n}$ be nonsingular and selfadjoint such that (2.2) is satisfied and let $A \in \mathbb{F}^{n \times n}$ such that $A^H = A^G$. Then there exists a nonsingular matrix $S \in \mathbb{F}^{n \times n}$ such that

$$S^{-1}AS = A_1 \oplus A_2, \quad S^HHS = J_1 \oplus J_2, \quad S^GS = \mu J_1 \oplus -\mu J_2,$$

where $A_1 \in \mathbb{F}^{(p+q) \times (p+q)}$, $A_2 \in \mathbb{F}^{(r+s) \times (r+s)}$ and $J_1 = I_p \oplus -I_q$, $J_2 = I_r \oplus -I_s$.

Proof. For the nonsingular matrix $S \in \mathbb{F}^{n \times n}$ from Lemma 2.4, the matrices $S^HHS$ and $S^GS$ take on the asserted form and $H^{-1}G = SFS^{-1}$ where $F = \mu I_{p+q} \oplus -\mu I_{r+s}$. According to the assumption $HAH^{-1} = GAG^{-1}$ we also have $F(S^{-1}AS) = S^{-1}(H^{-1}GA)S = S^{-1}(AH^{-1}G)S = (S^{-1}AS)F$, which is possible only if $S^{-1}AS$ has the asserted form. \hfill $\square$

If the matrix $A$ satisfies $A^H = A^G$ and, furthermore, admits an $H$-polar decomposition, then although

$$G^{-1}M^*U^*G = H^{-1}M^*U^*H = H^{-1}M^*HH^{-1}U^*H = MU^{-1}$$

or $M^*U^*G = GM$, it cannot be concluded that the matrix also admits a $G$- or a $(G, H)$-polar decomposition. However, the following statement holds.

Lemma 2.6 (F $\in \mathbb{R}$ or $F = \mathbb{C}$). Let $G, H, A, S \in \mathbb{F}^{n \times n}$ be as in Lemma 2.5. Then $A$ admits a $(G, H)$-polar decomposition if and only if $A_1$ admits a $J_1$-polar decomposition and $A_2$ admits a $J_2$-polar decomposition. Moreover, such a decomposition is $H$-semidefinite if and only if both $J_k$-polar decompositions are semidefinite.
Proof. Let $A = UM$ be a $(G,H)$-polar decomposition. Then $U^H = U^G$ and $M^H = M^G$ imply $S^{-1}US = U_1 \oplus U_2$ and $S^{-1}MS = M_1 \oplus M_2$, where the blocks $A_k, J_k, U_k, M_k$ have the same size ($k = 1, 2$). A simple calculation shows that $U_kM_k$ is a $J_k$-polar decomposition of $A_k$.

If conversely $A_1 = U_1M_1$ and $A_2 = U_2M_2$ are given $J_1$- and $J_2$-polar decompositions, then these are also $(\mu J_1)$- and $(-\mu J_2)$-polar decompositions and therefore $A = UM$ with $U = S(U_1 \oplus U_2)S^{-1}$ and $M = S(M_1 \oplus M_2)S^{-1}$ is a $(G,H)$-polar decomposition.

The second part of the assertion follows from the fact that $HM \geq 0$ if and only if $J_kM_k \succeq 0$ for $k = 1, 2$. $\square$

A useful application of this lemma is the next result which ensures the existence of a $(G,H)$-polar decomposition in an important particular case.

**Lemma 2.7** ($F = \mathbb{R}$ or $F = \mathbb{C}$). Let $G, H \in F^{n \times n}$ be nonsingular and selfadjoint such that (2.2) is satisfied and let $A \in F^{n \times n}$ such that $A^H = A^G$. If $A = UM$ is an $H$-polar decomposition with $\sigma(M) \subset \mathbb{C}^+ = \{z \in \mathbb{C} | \text{Re}(z) > 0\}$, then this is also a $G$-polar decomposition.

**Proof.** Let $S$ be as in Lemma 2.5. Then from $\sigma(S^{-1}A^HAS) = \sigma(A^H A) = \sigma(M^2) \subset \mathbb{C} \setminus (-\infty, 0]$ and

$$(S^{-1}H^{-1}S^{-*})(S^*A^*S^{-*})(S^*HS)(S^{-1}AS) = \bigoplus_{k=1}^2 J_k^{-1}A_k^*J_kA_k$$

it follows that $\sigma(A_k^H A_k) \subset \mathbb{C} \setminus (-\infty, 0]$ for $k = 1, 2$. Thus, according to Proposition 2.1, both blocks $A_k$ admit a $J_k$-polar decomposition $U_kM_k$ with $\sigma(M_k) \subset \mathbb{C}^+$. Moreover, Lemma 2.6 implies that

$$A = \tilde{U}\tilde{M}$$

with $\tilde{U} = S(U_1 \oplus U_2)S^{-1}$ and $\tilde{M} = S(M_1 \oplus M_2)S^{-1}$

is a $(G,H)$-polar decomposition with $\sigma(M) \subset \mathbb{C}^+$. On the other hand, according to [15, Section 4], there exists one and only one matrix $M$ for which $A^H A = M^2$ and $\sigma(M) \subset \mathbb{C}^+$, so that $\tilde{M} = M$ and thus also $U = \tilde{U}$ must be true. $\square$

In conclusion of this preparatory section, the statements of the lemmas $^3$ will be explained with the help of three examples.

**Example 2.8.** Let $H = I_p \oplus I_r$ and $G = I_p \oplus -I_r$. Then a matrix $A \in F^{(p+r) \times (p+r)}$ for which $A^H = A^G$, according to Lemma 2.5, takes on the form $A = A_1 \oplus A_2$, where $A_1 \in F^{p \times p}$ and $A_2 \in F^{r \times r}$. Let

$^3$ More general results on $(G,H)$-polar decompositions have been found for the case in which $\rho H - G$ is a non-defective Hermitian pencil. However, their presentation involves the discussion of an algebraic Riccati equation which is beyond the scope of this paper.
A_1 = P_1 \Sigma_1 Q_1^* \quad \text{and} \quad A_2 = P_2 \Sigma_2 Q_2^*. \\

be singular value decompositions and let \\
U = P_1 Q_1^* \oplus P_2 Q_2^* \quad \text{and} \quad M = Q_1 \Sigma_1 Q_1^* \oplus Q_2 \Sigma_2 Q_2^*. \\

Then A = UM is an H-semidefinite \((G, H)\)-polar decomposition.

**Example 2.9.** Let \(\alpha, \beta, \mu \in \mathbb{R}\) with \(\mu \neq 0\) and let \(H = \text{diag}(1, -1, 1, -1)\) and \(G = \mu \text{diag}(1, -1, -1, 1)\). The matrix \(A_1 = \begin{bmatrix} 0 & \beta \\ \alpha & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix}\) satisfies \(A_1^H A_1 = A_1^G A_1 = \text{diag}(-\alpha^2, -\beta^2, -\beta^2, -\alpha^2)\) and admits the \(H\)-polar decomposition

\[
A_1 = U_1 M_1 \quad \text{with} \quad U_1 = \begin{bmatrix} -i & 0 \\ -i & 0 \end{bmatrix}, \quad M_1 = \begin{bmatrix} 0 & i \alpha \\ i \alpha & 0 \end{bmatrix}.
\]

But it is not a \(G\)-polar decomposition because \(U_1^* G U_1 = -G\) and \(M_1^* G = -G M_1\).

In fact when \(\alpha \neq \beta\), the matrix pair \((A_1^H A_1, G)\), which is already in canonical form, does not satisfy the condition (2.1). So \(A_1\) does not have any \(G\)-polar decompositions in this case. The matrix \(A_2 = \begin{bmatrix} 0 & \beta \\ \alpha & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & \beta \\ \alpha & 0 \end{bmatrix}\) satisfies \(A_2^H A_2 = A_2^G A_2 = \text{diag}(-\alpha^2, -\beta^2, -\alpha^2, -\beta^2)\) and admits the \(G\)-polar decomposition

\[
A_2 = U_2 M_2 \quad \text{with} \quad U_2 = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} i \alpha & 0 \\ 0 & i \beta \end{bmatrix}.
\]

But it is not an \(H\)-polar decomposition because \(U_2^* H U_2 = -H\) and \(M_2^* H = -H M_2\).

Again when \(\alpha \neq \beta\), the matrix pair \((A_2^H A_2, H)\), which is already in canonical form, does not satisfy the condition (2.1). So \(A_2\) does not have any \(H\)-polar decompositions in this case. But if \(\alpha = \beta\), then \(A = A_1 = A_2\) admits the \((G, H)\)-polar decomposition

\[
A = UM \quad \text{with} \quad U = \begin{bmatrix} -i & 0 \\ -i & 0 \end{bmatrix}, \quad M = \begin{bmatrix} i \alpha & 0 \\ 0 & i \beta \end{bmatrix}.
\]

which evidently satisfies Lemma 2.6.

**Example 2.10.** Let \(G, H\) be matrices with (2.2) and let \(A\) be a matrix with \(A^H = A^G\). If \(H\) is positive definite and \(A\) nonsingular, then there exists a definite \(H\)-polar
decomposition which, according to Lemma 2.7, is a G-polar decomposition too. However, if A is singular or H is indefinite, this may not be always true. Consider the (semi)define H-polar decompositions

\[ H_1 = \text{diag}(1, 1, 1), \quad G_1 = \text{diag}(1, 1, -1), \quad x \in \mathbb{R}, \]
\[ A_1 = \begin{bmatrix} \cos(x) & 0 & 0 \\ \sin(x) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \cos(x) & 0 & -\sin(x) \\ \sin(x) & 0 & \cos(x) \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = U_1 M_1, \]

where \( \sigma(H_1 M_1) = \{0, 1\} \) and

\[ H_2 = \text{diag}(1, -1), \quad G_2 = \text{diag}(1, 1), \quad a > |b| > 0, \quad u = b/a, \]
\[ A_2 = \begin{bmatrix} -\sqrt{a^2 - b^2} & 0 \\ 0 & \sqrt{a^2 - b^2} \end{bmatrix} = \frac{1}{\sqrt{1 - u^2}} \begin{bmatrix} 1 & u \\ u & 1 \end{bmatrix} \begin{bmatrix} a & b \\ -b & -a \end{bmatrix} = U_2 M_2, \]

where \( \sigma(H_2 M_2) = \{a \pm b\} \). Here \( U_1^* G_1 U_1 = \text{diag}(1, -1, 1) \neq G_1 \) and neither \( U_2 \) is orthogonal nor \( M_2 \) symmetric, so that both factorisations are not G-polar decompositions. In contrast to this, the “blockwise” (semi)define H-polar decompositions

\[ A_1 = \begin{bmatrix} \cos(x) \\ -\sin(x) \\ \sin(x) \end{bmatrix} \oplus 1 \]
\[ \oplus 1 \]
\[ G_1 \]
\[ \quad M_1 \quad \]
\[ \quad A_2 = (-I_2)(-A_2), \]

according to Lemma 2.6, are also G-polar decompositions.

With this background on \( H \)- and \((G, H)\)-polar decompositions we are now able to investigate the problems stated in the introduction, starting with the determination of vectors \( x_1, \ldots, x_N \) and an indefinite scalar product \[ \langle \cdot, \cdot \rangle \] such that \[ \langle x_k - x_l, x_k - x_l \rangle = q_{kl} \] for given real numbers \( q_{kl} \) (\( 1 \leq k, l \leq N \)).

### 3. Construction of vectors from values of a quadratic form

The construction of vectors from given values of a quadratic form presented in this section is a generalisation of the work [21] for complex vector spaces and indefinite scalar products.

Let \( F = \mathbb{R} \) or \( F = \mathbb{C} \) and let \[ \langle \cdot, \cdot \rangle \] be an indefinite scalar product in \( F^n \) with the underlying nonsingular symmetric or Hermitian matrix \( H \in F^{n \times n} \). Then for arbitrary vectors \( x, y \in F^n \) in the case \( F = \mathbb{R} \) it is true that

\[ \langle x, y \rangle = \frac{1}{2} \langle [x, x] + [y, y] - [x - y, x - y] \rangle \]  \hspace{1cm} (3.1a)

and in the case \( F = \mathbb{C} \) we have

\[ \text{Re}\langle x, y \rangle = \frac{1}{2} \langle [x, x] + [y, y] - [x - y, x - y] \rangle \]
\[ \quad = \frac{1}{2} \langle [x, x] + [y, y] - [iy - ix, iy - ix] \rangle, \]  \hspace{1cm} (3.1b)
The scalar products of the vectors can be expressed in terms of the quadratic form \( \Phi(x) = [x, x] \).

Now let \( n \leq N \) vectors \( x_1, \ldots, x_N \in \mathbb{F}^n \) be given and let \( X = [x_1 \ldots x_N] \in \mathbb{F}^{n \times N} \) be a matrix whose columns are these vectors. Then

\[
W = X^*HX
\]

is the Gramian matrix of the \( x_k \). Therefore, if \( \text{span}\{x_1, \ldots, x_N\} = \mathbb{F}^n \), then the number of positive and negative eigenvalues of \( H \) and \( W \) are equal, and furthermore the eigenvalue 0 appears in \( \sigma(W) \) with the multiplicity \( N - n \) (Sylvester’s law of inertia, [7, Chapter IX, §2]). Moreover, the elements \( w_{kl} = [x_l, x_k] \) of the matrix \( W \) according to (3.1) can be expressed in the form

\[
\begin{align*}
w_{kl} &= \frac{1}{2}(\rho_k + \rho_l - \sigma_{kl}) & \text{if } \mathbb{F} = \mathbb{R} \text{ or} \\
w_{kl} &= \frac{1}{2}(\rho_k + \rho_l - \sigma_{kl}) + \frac{i}{2}(\rho_k + \rho_l - \tau_{kl}) & \text{if } \mathbb{F} = \mathbb{C},
\end{align*}
\]

where

\[
\begin{align*}
\rho_k &= [x_k, x_k], \quad \sigma_{kl} = [x_l - x_k, x_l - x_k], \quad \tau_{kl} = [x_l - ix_k, x_l - ix_k], \\
\text{with } \rho_k, \sigma_{kl}, \tau_{kl} &\in \mathbb{R}, \quad \sigma_{kl} = \sigma_{lk}, \quad \sigma_{kk} = 0, \quad \tau_{kl} + \tau_{lk} = 2(\rho_k + \rho_l)
\end{align*}
\]

for \( 1 \leq k, l \leq N \).

Conversely, let the real numbers \( \rho_k, \sigma_{kl}, \tau_{kl} \) with (3.4) be given, and let the elements of a matrix \( W \) be defined by (3.2). Then this matrix is symmetric or Hermitian, respectively, and can therefore be written in the form

\[
W = R^*AR.
\]

Here \( A \) is a diagonal matrix of the real eigenvalues \( \lambda_1, \ldots, \lambda_N \) of \( W \) and \( R = [r_1 \ldots r_N] \) is a matrix whose columns form a basis of \( \mathbb{F}^N \) consisting of orthonormalised eigenvectors. Now if \( p \) is the number of positive and \( n - p \) is the number of negative eigenvalues and if it is assumed that

\[
\lambda_1, \ldots, \lambda_p > 0, \quad \lambda_{p+1}, \ldots, \lambda_n < 0 \quad \text{and} \quad \lambda_{n+1} = \cdots = \lambda_N = 0,
\]

then the matrices defined by

\[
A_1 = \text{diag}(\lambda_1, \ldots, \lambda_p, \lambda_{p+1}, \ldots, \lambda_n) \quad \text{and} \quad R_1 = [r_1 \ldots r_n]
\]

satisfy \( W = R_1 A_1 R_1^* \) too. Consequently, if we set

\[
\Sigma = \text{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_p}, \sqrt{-\lambda_{p+1}}, \ldots, \sqrt{-\lambda_n}) \quad \text{and} \quad H_p = I_p \oplus -I_{n-p},
\]

the following properties hold.
then the matrix
\[ X = \Sigma^* R_1^* \in \mathbb{F}^{n \times N} \]
fulfills on the one hand \( \text{rank} X = n \) and on the other hand \( X^* H_\sigma X = R_1 \Sigma H_\sigma \Sigma^* R_1^* = R_1 A R_1^* = W \). Therefore the columns \( x_1, \ldots, x_N \in \mathbb{F}^n \) of \( X \) constitute a spanning set (or system of generators) for \( \mathbb{F}^n \), and for the indefinite scalar product defined by \( [x, y]_w = (H_\sigma x, y) \) it is true that \( w_{kl} = [x_l, x_k]_w \). This means that also
\[
[x_l, x_k]_w = w_{kk} = \rho_k, \\
[x_l - x_k, x_l - x_k]_w = w_{kk} + w_{ll} - w_{kl} - w_{lk} = \sigma_{kl}, \\
[x_l - ix_k, x_l - ix_k]_w = w_{kk} + w_{ll} + iw_{kl} - iw_{lk} = \tau_{kl}
\]
and this determinant is non-negative if and only if \( w_{kk} \leq \sigma_{kl} \) and \( \sqrt{\rho_1 - \sigma_{12}} \leq \sqrt{\rho_2} - \sigma_{22} \).

Theorem 3.1 (Construction of vectors). Let \( \mathbb{F} = \mathbb{C} \) and let \( \rho_k, \sigma_{kl} \) be real numbers such that \( \sigma_{kl} = \sigma_{lk} \) and \( \sigma_{kk} = 0 \) for all \( k, l \) in \( \{1, \ldots, N\} \). Furthermore, for the case \( \mathbb{F} = \mathbb{C} \) let \( \tau_{kl} \) be real numbers such that \( \tau_{kl} + \tau_{lk} = 2(\rho_k + \rho_l) \) for all \( k, l \) in \( \{1, \ldots, N\} \). Then the following statements are equivalent:

(i) There exist vectors \( x_1, \ldots, x_N \in \mathbb{F}^n \) constituting a spanning set for \( \mathbb{F}^n \), for which \( [x_k, x_k]_w = \rho_k \) as well as \( [x_l - x_k, x_l - x_k]_w = \sigma_{kl} \), and in the case \( \mathbb{F} = \mathbb{C} \) also \( [x_l - ix_k, x_l - ix_k]_w = \tau_{kl} \) is satisfied. Thereby \([\cdot, \cdot]_w\) is an indefinite scalar product in \( \mathbb{F}^n \) with underlying nonsingular symmetric or Hermitian matrix \( H \in \mathbb{F}^{n \times n} \) which has \( p \) positive eigenvalues.

(ii) The symmetric or Hermitian matrix \( W \in \mathbb{F}^{N \times N} \) whose elements \( w_{kl} \) are defined by (3.2) has \( p \) positive and \( n - p \) negative eigenvalues, and the eigenvalue \( 0 \) appears with multiplicity \( N - n \).

For the case of a Euclidean or unitary space we immediately obtain the following corollary in which \( \| \cdot \| \) denotes the Euclidean norm.

Corollary 3.2. Let \( \rho_k, \sigma_{kl}, \tau_{kl} \geq 0 \) be as in Theorem 3.1. Then there exist vectors \( x_k \) such that \( \|x_k\| = \sqrt{\rho_k}, \|x_l - x_k\| = \sqrt{\sigma_{kl}}, \) and in the case \( \mathbb{F} = \mathbb{C} \) also \( \|x_l - ix_k\| = \sqrt{\tau_{kl}} \) if and only if the matrix \( W \) is positive semidefinite.

Let \( \mathbb{F} = \mathbb{R}, N = 2 \) and \( \rho_1, \rho_2, \sigma_{12} \geq 0 \). Then
\[
\det W = \left( \frac{1}{2} \rho_1 \rho_2 + \rho_1 \sigma_{12} + \rho_2 \sigma_{12} \right) - \frac{1}{4} \left( \rho_1^2 + \rho_2^2 + \sigma_{12}^2 \right) \\
= \frac{1}{4} \left( \sigma_{12} - (\sqrt{\rho_1} - \sqrt{\rho_2})^2 \right) \left( (\sqrt{\rho_1} + \sqrt{\rho_2})^2 - \sigma_{12} \right)
\]
and this determinant is non-negative if and only if
\[
|\sqrt{\rho_1} - \sqrt{\rho_2}| \leq \sqrt{\sigma_{12}} \quad \text{and} \quad \sqrt{\sigma_{12}} \leq \sqrt{\rho_1} + \sqrt{\rho_2}.
\]
But this is just the triangle inequality, so that Corollary 3.2 gives a generalisation of this essential property of Euclidean geometry.

In addition to these investigations concerning the geometrical properties of the vectors $x_k$, the consideration of their physical properties provides some useful information for the application of Theorem 3.1 in MDS.

**Remark 3.3 (Tensor of inertia).** On interpreting the vectors $x_k = (x_\alpha^k)$ constructed in Theorem 3.1 as the locations of point objects of mass 1, the matrix

$$T = XX^*$$

with

$$T_{\alpha\beta}^{\sigma\rho} = \sum_{k=1}^{N} x_\alpha^k x_\beta^k$$

for $1 \leq \alpha, \beta \leq n$

gives their (contravariant) tensor of inertia in the sense of Hermann Weyl [20, §6]. Here $T = \Sigma^* R \Sigma; \Sigma = \Sigma^2 = \text{diag}(|\lambda_1|, \ldots, |\lambda_n|)$ is a diagonal matrix, so that the axes of the coordinate system are also the inertial axes (principal axes) of the constellation. Moreover, the absolute values of the eigenvalues give the associated (contravariant) moments of inertia. From the viewpoint of MDS this means that the coordinates of the vectors can be interpreted, as usual, as the ratings of uncorrelated factors [1, Section 7.10].

**Remark 3.4 (Centroid).** Let $x_1, \ldots, x_N \in \mathbb{R}^n$ be real vectors whose centroid lies at the coordinates’ origin, i.e. $\sum x_k = 0$, and let $\Phi(x) = [x, x]$. Then the scalar products satisfy

$$[x_i, x_k] = \frac{1}{2} \left( \frac{1}{N} \sum_{j=1}^{N} \Phi(x_k - x_j) + \frac{1}{N} \sum_{i=1}^{N} \Phi(x_i - x_i) 
- \Phi(x_k - x_i) - \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \Phi(x_i - x_j) \right),$$

as can easily be verified [19]. Conversely, let the real numbers $\sigma_{kl} = \sigma_{lk}, \sigma_{kk} = 0, 1 \leq k, l \leq N$ be given. Then

$$w_{kl} = \frac{1}{2} \left( \frac{1}{N} \sum_{j} \sigma_{kj} + \frac{1}{N} \sum_{i} \sigma_{il} - \sigma_{kl} - \frac{1}{N^2} \sum_{i} \sum_{j} \sigma_{ij} \right)$$

defines the elements of a symmetric matrix $W$ whose row and column sums vanish. Using the method of Theorem 3.1 again vectors $x_k$ and an indefinite scalar product can be constructed such that $w_{kl} = [x_l, x_k]$. But now the centroid of these vectors lies at the origin. An analogous construction also applies in the complex case, but the conditions that must be assumed for the values $\tau_{kl}$ are rather complicated there.

---

4 Weyl’s definition slightly differs from the definition given in the textbooks of classical physics. Nevertheless, it is more reasonable when considering the rotational motion in $n$-dimensional spaces.
4. Solution of the $H$-isometric Procrustes problem

Let $x_1, \ldots, x_N \in \mathbb{F}^n$ be the vectors and let $[.,.] = (H.\ldots)$ be the indefinite scalar product constructed from given scalars $\rho_k, \sigma_{kl}, \tau_{kl}$ according to Theorem 3.1, so that (3.3) holds. For every $H$-isometry $U \in \mathbb{F}^{n \times n}$ it then follows that

$$[Ux_k, Ux_k] = [x_k, x_k] = w_{kl},$$

which can also be expressed in matrix equation form

$$X^*U^*HU = X^*HX = W.$$ 

Thus the columns $x'_k = Ux_k$ contained in the matrix $X' = UX$ satisfy (3.3), too. Now assume that $x_1, \ldots, x_N$ and $y_1, \ldots, y_N$ are the vectors constructed from two measurements of a quadratic form. Then on comparing the constellations the question arises, what part of the observed differences is due to different positions in space, and what part is due to actual differences in the inner structure of the constellations. Expressed mathematically, the task is to determine an $H$-isometry $U \in \mathbb{F}^{n \times n}$ which solves the optimisation problem

$$f(U) = \sum_{k=1}^{N} [Ux_k - y_k, Ux_k - y_k] \rightarrow \begin{cases} 
\min, & \text{if } H > 0 \\
\max, & \text{if } H < 0 \\
\min / \max, & \text{otherwise}
\end{cases} \quad (4.1a)$$

$$h(U) = U^*HU - H = 0.$$ 

The sum of scalar products arising therein can be expressed in the form of a trace, so that an alternative expression with

$$f(U) = \text{tr}((UX - Y)^*H(UX - Y)) \quad (4.1b)$$

is given, where as above $X = [x_1 \ldots x_N]$ and $Y = [y_1 \ldots y_N]$. Moreover, $H < 0$ ($H > 0$) stands for a positive (negative) definite matrix $H$ and the symbol “min / max” stands for a particular saddle point, which will be explained more precisely below. Within the scope of Euclidean vector spaces a solution of this problem was found in [17] where it was called the orthogonal Procrustes problem ($\mathbb{F} = \mathbb{R}, H = I$). In the present context of indefinite scalar products it is furthermore called the $H$-orthogonal or $H$-unitary Procrustes problem.

The fact, that the addends in (4.1) can be positive as well as negative, whenever $H$ is indefinite, causes severe difficulties. On first sight one may thus get the idea to avoid these difficulties by minimising one of the non-negative functions

$$f_1(U) = f(U)^2 \geq 0 \quad \text{or} \quad f_2(U) = \sum_k [Ux_k - y_k, Ux_k - y_k]^2 \geq 0.$$ 

But the example $H = \text{diag}(1, -1), Ux_k = (\xi, \xi)^T, y_k = (\eta, \eta)^T$, i.e.

$$[Ux_k - y_k, Ux_k - y_k] = |\xi - \eta|^2 - |\xi - \eta|^2 = 0,$$

shows an addend which neither in $f_1$ nor in $f_2$ makes a contribution to the result although $|\xi - \eta|$ may be arbitrarily large. However, the intention of the optimisation is
to converge the constellations in the sense of an optimum congruence which means, that the coordinate differences should become small. A first possibility to reach this goal is to measure the differences with a definite scalar product, e.g. \( \| Ux_k - y_k \|_2 \). This approach will be discussed in the next section. A further possibility is not to look for a minimum or maximum of the function \( f \), but to determine a particular saddle point “min /max” where the coordinate differences are small. This is the subject of the following investigations.

Considering the case \( F = \mathbb{R} \) first and introducing a matrix of the (unknown) Lagrange multipliers \( L \in \mathbb{R}^{n \times n} \), the constraints can be stated in the form

\[
h_L(U) = \text{tr}[L(U^*HU - H)]
\]

and the necessary first order condition for solving the problem is

\[
\frac{\partial}{\partial U}(f + h_L) = 0.
\]

Differentiation of the trace [5] gives

\[
\frac{\partial f}{\partial U} = 2HUXX^* - 2HYX^* \quad \text{and} \quad \frac{\partial h_L}{\partial U} = HU(L + L^*),
\]

so that \( U \) must satisfy the equation

\[
UXX^*H + UAH = YX^*H \quad \text{with} \quad \Lambda = \frac{L + L^*}{2} = \Lambda^*.
\]

(4.2)

Now defining \( M = (XX^* + \Lambda)H \), the necessary condition becomes

\[
A = UM \quad \text{with} \quad A = YX^*H \quad \text{and} \quad U^*HU = H, \quad M^*H = HM.
\]

(4.3)

Thus, if a solution of the problem exists, it can be determined by an \( H \)-polar decomposition of the matrix \( A \). (The question which of the \( H \)-isometries contained in such an \( H \)-polar decomposition actually are solutions of the problem will be discussed after the complex case is complete.)

In the case \( F = \mathbb{C} \) the complex derivatives of \( f \) and \( h_L \) do not exist. However, the necessity for (4.3) can be shown by determining the real derivatives. For this, let the real and imaginary part of the matrix \( A \in \mathbb{C}^{m \times n} \) be denoted by \( A_1 \) and \( A_2 \), respectively. Then the well-known linear map

\[
T : \mathbb{C}^{m \times n} \rightarrow \mathbb{R}^{2m \times 2n},
\]

\[
T(A) = Q_{2m}^* \begin{bmatrix} A & 0 \\ 0 & \overline{A} \end{bmatrix} Q_{2n} = \begin{bmatrix} A_1 & A_2 \\ -A_2 & A_1 \end{bmatrix}, \quad \text{where} \quad Q_{2n} = \frac{\sqrt{2}}{2} \begin{bmatrix} I_n & -iI_n \\ iI_n & I_n \end{bmatrix},
\]

allows the real representation \( A^\wedge = T(A) \) of \( A \). Moreover, for every Hermitian matrix \( A \) it is true that \( 2\text{tr}(A) = \text{tr}(A^\wedge) \) which follows from the unitarity of \( Q_{2n} \). Therefore, the objective function can be represented as

\[
2f(U) = f(U^\wedge) = \text{tr}[(U^\wedge X^\wedge - Y^\wedge)^T H^\wedge(U^\wedge X^\wedge - Y^\wedge)]
\]

having the real derivatives

\[
\frac{\partial f(U^\wedge)}{\partial U^\wedge} = \begin{bmatrix} \frac{\partial f}{\partial U_1} \\ \frac{\partial f}{\partial U_2} \end{bmatrix} = 2H^\wedge U^\wedge X^\wedge(T) - 2H^\wedge Y^\wedge(T).
The transformation of the constraints
\[ h_1(U) = \text{Re}(U^*HU - H) = 0 \quad \text{and} \quad h_2(U) = \text{Im}(U^*HU - H) = 0 \]
is more complicated. Introducing Lagrange multipliers \( L_1, L_2 \in \mathbb{R}^{n \times n} \) and using \( H = H_1 + iH_2, U = U_1 + iU_2 \) we obtain
\[
\begin{align*}
    h_{L,1}(U) &= \text{tr}[L_1(U_1^TH_1U_1 - U_1^TH_2U_2 + U_2^TH_1U_2 + U_2^TH_2U_1 - H_1)], \\
    h_{L,2}(U) &= \text{tr}[L_2(U_1^TH_1U_1 - U_1^TH_2U_2 - U_2^TH_1U_2 + U_2^TH_2U_1 - H_2)],
\end{align*}
\]
from which it follows that
\[
\begin{align*}
    \frac{\partial h_{L,1}}{\partial U_1} &= (H_1U_1 - H_2U_2)(L_1 + L_1^T), \\
    \frac{\partial h_{L,1}}{\partial U_2} &= (H_1U_2 + H_2U_1)(L_1 + L_1^T), \\
    \frac{\partial h_{L,2}}{\partial U_1} &= (H_2U_1 + H_1U_2)(L_2 - L_2^T), \\
    \frac{\partial h_{L,2}}{\partial U_2} &= (H_2U_2 - H_1U_1)(L_2 - L_2^T),
\end{align*}
\]
where \( H_1 = H_1^T \) and \( H_2 = -H_2^T \) must be taken into account. Now setting
\[
    A_1 = \frac{L_1 + L_1^T}{2} = A_1^*, \quad A_2 = \frac{L_2 - L_2^T}{2} = -A_2^* \quad \text{and} \quad \Lambda = A_1 + iA_2 = \Lambda^*,
\]
it can be verified that
\[
\begin{align*}
    \frac{\partial h_L(U^\wedge)}{\partial U^\wedge} &= \begin{pmatrix}
        \frac{\partial h_{L,1}(U^\wedge)}{\partial U_1} & \frac{\partial h_{L,2}(U^\wedge)}{\partial U_2} \\
        \frac{\partial h_{L,1}(U^\wedge)}{\partial U_2} & \frac{\partial h_{L,2}(U^\wedge)}{\partial U_1}
    \end{pmatrix} = 2H^*U^\wedge(\overline{\Lambda})^\wedge.
\end{align*}
\]
Consequently, the necessary first order conditions for an optimum
\[
\frac{\partial}{\partial U_1}(f + h_{L,1} + h_{L,2}) = 0 \quad \text{and} \quad \frac{\partial}{\partial U_2}(f + h_{L,1} + h_{L,2}) = 0
\]
can be stated as
\[
\frac{\partial f(U^\wedge)}{\partial U^\wedge} + \frac{\partial h_L(U^\wedge)}{\partial U^\wedge} = 2(HUX^* - HYX^* + HU\overline{\Lambda})^\wedge = 0,
\]
showing that (4.2) and (4.3) must be satisfied in the complex case, too. (The conjugation is irrelevant since \( \Lambda \) may simply be renamed to \( \Lambda^* \).)

It remains to determine the particular \( H \)-polar decomposition (if existent) which leads to the optimum congruence. For this, let \( U^\Lambda \) be an \( H \)-polar decomposition of the matrix \( A = YX^*H \), and let
\[
\begin{align*}
    (R^{-1}\Lambda^{*1}A^R, R^*HR) &= (J, Z_J) \quad \text{and} \quad (S^{-1}MS^*, S^*HS) &= (K, Z_K)
\end{align*}
\]
be the canonical forms [6, Theorem I.3.3] of the pairs \( (\Lambda^{*1}A, H) = (M^2, H) \) and \( (M, H) \), respectively. Returning to the initial equation (4.1b), we find that
\[
\begin{align*}
    f(U) &= \text{tr}[U(X - Y)^*H(UX - Y)] \\
    &= \text{tr}(X^*U^*HUX - X^*U^*HY - Y^*HUX + Y^*HY)
\end{align*}
\]
\[ \begin{align*}
&= \text{tr}(X^*HX) + \text{tr}(Y^*HY) - 2\text{Re tr}[YX^*H(\mathbf{H}^{-1}U^*\mathbf{H})] \\
&= \tau - 2\text{Re tr}(AU^{-1}) = \tau - 2\text{Re tr}(UMU^{-1}) \\
&= \tau - 2\text{Re tr}(\mathbf{SKS}^{-1}) = \tau - 2\text{Re tr}(\mathbf{K}),
\end{align*} \]

where \( \tau = \text{tr}(X^*HX) + \text{tr}(Y^*HY) \). The optimum can be found from this equation by considering three cases:

Case (a): If \( \mathbf{H} \) is definite, then the canonical forms are of the kind
\[
(J_J, Z_J) = \left( \bigoplus_{j=1}^{k} \lambda_j \mathbf{I}_{p_j} \oplus \mathbf{0}_r, \bigoplus_{j=1}^{k} \delta \mathbf{I}_{p_j} \oplus \mathbf{0}_r \right),
\]
\[
(K_K, Z_K) = \left( \bigoplus_{j=1}^{k} \sqrt{\lambda_j} \mathbf{I}_{p_j} \oplus \mathbf{0}_r, \bigoplus_{j=1}^{k} \varepsilon \mathbf{I}_{p_j} \oplus \mathbf{0}_r \right).
\]

Here \( \lambda_j > 0, \Sigma_{p_j} = \text{diag}(\pm 1) \) for \( 1 \leq j \leq k \) and \( \varepsilon = +1 \) if \( \mathbf{H} > 0, \varepsilon = -1 \) if \( \mathbf{H} < 0 \). In the case \( \mathbf{H} > 0 \) the value \( f(U) \) takes its minimum, when \( \Sigma_{p_j} = +\mathbf{I}_{p_j} \) is chosen and in the case \( \mathbf{H} < 0 \) the value \( f(U) \) takes its maximum, when \( \Sigma_{p_j} = -\mathbf{I}_{p_j} \) is chosen. This means that in both cases
\[ Z_K \mathbf{K} = \bigoplus_{j=1}^{k} \sqrt{\lambda_j} \mathbf{I}_{p_j} \oplus \mathbf{0}_r \geq 0 \]

and thus \( \mathbf{HM} \geq 0 \), so that the wanted result is obtained via a semidefinite \( \mathbf{H} \)-polar decomposition of \( \mathbf{A} \). In particular, if \( \mathbf{H} = \mathbf{I} \), then the solution is determined by an ordinary polar decomposition where \( \mathbf{U}^* = \mathbf{U}^{-1} \) and \( \mathbf{M}^* = \mathbf{M} \) is positive semidefinite.

Case (b): If \( \mathbf{H} \) is indefinite and if \( \mathbf{A} \) admits a semidefinite \( \mathbf{H} \)-polar decomposition, then the following relationships exists between the canonical forms
\[
\mathbf{J} = \bigoplus_{j=1}^{k} \begin{bmatrix} \lambda_j & 0_r \\ 0_{r+s} & \lambda_j \end{bmatrix} \oplus \begin{bmatrix} \mathbf{0}_{r+s} \\ \mathbf{0}_{r+s} \end{bmatrix},
\]
\[
\mathbf{Z}_J = \bigoplus_{j=1}^{k} \begin{bmatrix} \mathbf{1}_{p_j} \\ -\mathbf{I}_{q_j} \end{bmatrix} \oplus \begin{bmatrix} \mathbf{1}_{r+s} \\ -\mathbf{1}_{r+s} \end{bmatrix}.
\]
\[
\mathbf{K} = \bigoplus_{j=1}^{k} \begin{bmatrix} \sqrt{\lambda_j} \Sigma_{p_j} \\ \sqrt{\lambda_j} \Sigma_{q_j} \end{bmatrix} \oplus \begin{bmatrix} \mathbf{0}_r \\ \mathbf{0}_r \end{bmatrix} \oplus \begin{bmatrix} \mathbf{0}_r \\ \mathbf{0}_r \end{bmatrix},
\]
\[
\mathbf{Z}_K = \bigoplus_{j=1}^{k} \begin{bmatrix} \mathbf{1}_{p_j} \\ -\mathbf{I}_{q_j} \end{bmatrix} \oplus \begin{bmatrix} \mathbf{1}_r \\ -\mathbf{1}_r \end{bmatrix} \oplus \begin{bmatrix} \mathbf{1}_r \\ -\mathbf{1}_r \end{bmatrix},
\]

(4.4a)
where \( \lambda_j > 0 \) for \( 1 \leq j \leq k \) [3, Theorem 5.3]. If in this case \( \sum p_j = I_p, \sum q_j = -I_q \) and \( \Sigma = I_r \) is chosen, then again

\[
Z_K K = \bigoplus_{j=1}^{k} \sqrt{\lambda_j} I_{p_j+q_j} \oplus 0_r, \oplus 0_s + t \succeq 0.
\]

(4.4b)

By this choice the contributions to \( f(U) \) take on their minimum along the positive space dimensions and their maximum along the negative space dimensions. This is what is meant by “min/max” in (4.1a). Moreover, the resulting coordinate differences are “small” which can be seen in the following way: Let \( \mathbf{X}' = UX \). Then

\[
Y(X)'^* = YX^*U^* = UMH^{-1}U^*
\]

\[
= U(SKS^{-1})(SZ_KS^*)U^*
\]

\[
= (US)KZ_K(US)^*
\]

is positive semidefinite since \( Z_K(Z_K K)Z_K = KZ_K \) is. Hence, the orthogonal or unitary Procrustes problem

\[
\varphi(T) = \text{tr}[(TX' - Y)(TX' - Y)] \rightarrow \min \quad \text{with} \quad T^*T = I,
\]

(4.5)

whose solution, according to case (a), is determined by an ordinary polar decomposition

\[
TM' = Y(X)'^* \quad \text{with} \quad M' = (M')^* \succeq 0
\]

is solved for \( T = I \). In other words, the coordinate differences \( \sum \|x_k' - y_k\|^2 \) obtained with the “min/max” solution \( x_k' = UX_k \) are at minimum with respect to an orthogonal or unitary transformation in the sense of problem (4.5). This is exactly what one would expect of a transformation to an optimum congruence.

Case (c): If \( H \) is indefinite and if \( A \) admits an \( H \)-polar decomposition but not a semidefinite \( H \)-polar decomposition, then by definition \( Z_K K \) and thus also \( KZ_K \) cannot be positive semidefinite. Therefore, there always exists a solution \( T_0 \) of the problem (4.5) for which \( \varphi(T_0) < \varphi(I) \). Hence, the wanted result of an optimum congruence of the constellations \( \mathbf{X}' \) and \( \mathbf{Y} \) cannot be achieved in this case.

This investigation shows that an \( H \)-isometry for which \( \mathbf{X}' = UX \) and \( \mathbf{Y} \) are at optimum congruence can only exist if \( A \) admits a semidefinite \( H \)-polar decomposition. Conversely, let \( \mathbf{X}' \) and \( \mathbf{Y} \) be matrices which are at optimum congruence, i.e. for which \( Y(X)'^* \) is positive semidefinite and selfadjoint. Moreover, let \( U \) be an \( H \)-isometry and let \( X = U^{-1}X' \). Then \( A = YX^*H = Y(X)'^*HU \) admits the semidefinite \( H \)-polar decomposition \( A = UM \) where \( M = U^{-1}Y(X)'^*HU \) is \( H \)-nonnegative. All in all, we thus have found the following result.

Theorem 4.1 (Solution of the \( H \)-isometric Procrustes problem). A solution of the \( H \)-orthogonal or \( H \)-unitary Procrustes problem (4.1) exists if and only if the matrix \( A = YX^*H \) admits a semidefinite \( H \)-polar decomposition. In this case the \( H \)-isometry \( U \) contained in such a decomposition \( A = UM \) optimises the function \( f \). Moreover,
$X' = UX$ and $Y$ are at optimum congruence in the sense, that the orthogonal or unitary Procrustes problem (4.5) is solved for $T = I$.

5. Solution of the $(G, H)$-isometric Procrustes problem

Whereas the $H$-isometric Procrustes problem can always be solved in the case of a definite matrix $H$, in the case of an indefinite matrix $H$ it is possible that no solution exists. But now let $G$ and $H$ be nonsingular selfadjoint matrices in $\mathbb{F}^{n \times n}$, a nd let the geometry within the tuples $(x_1, \ldots, x_N)$ and $(y_1, \ldots, y_N)$ be measured with the scalar product $[.,.]_G = (G, .)$, but the geometry between the tuples be measured with the scalar product $[.,.]_H = (H, .)$. Then the problem can be expressed, instead of (4.1), as

$$f(U) = \sum_{k=1}^{N} [Ux_k - y_k, Ux_k - y_k]_H \to \begin{cases} \min, & \text{if } H > 0 \\ \max, & \text{if } H < 0 \\ \min / \max, & \text{otherwise} \end{cases} (5.1a)$$

with $g(U) = U^*GU - G = 0$ and $h(U) = U^*HU - H = 0$

or in matrix notation

$$f(U) = \text{tr}[(UX - Y)^*H(UX - Y)], (5.1b)$$

which will be called the $(G, H)$-orthogonal or $(G, H)$-unitary Procrustes problem. If the vectors $x_k$ and $y_k$ result from a construction according to Theorem 3.1, the internal metric $G$ is fixed, but the external metric $H$ may be chosen within the scope of the “compatibility condition”

$$H^{-1}G = \mu^2G^{-1}H \quad \text{for some } \mu \in \mathbb{R} \setminus \{0\} (5.2)$$

which is characterised in Lemma 2.4. If this choice is made such that $H$ is positive definite, then a sum of non-negative distance squares is minimised. In this case a solution of (5.1) under the assumption (5.2) always exists which will be shown in the sequel. (An analogous statement holds for a negative definite matrix $H$.)

If again $L_G, L_H \in \mathbb{R}^{n \times n}$ are matrices of the (unknown) Lagrange multipliers and if the constraints in the case $\mathbb{F} = \mathbb{R}$ are stated in the form

$$g_L(U) = \text{tr}[L_G(U^*GU - G)] \quad \text{and} \quad h_L(U) = \text{tr}[L_H(U^*HU - H)],$$

then the necessary first order condition

$$\frac{\partial}{\partial U}(f + g_L + h_L) = 0$$

leads in the same way as above to the equation

$$\begin{align*}
GUA + HUB &= \tilde{C} \quad \text{with } \tilde{C} = HYX^* \\
A &= \frac{L_G + L_H}{2} = A^*, \quad B = XX^* + \frac{L_H + L_G}{2} = B^*
\end{align*} (5.3)$$
which is also valid in the case $F = \mathbb{C}$. Furthermore

$$GUG^{-1} = U^{-*} = HUH^{-1},$$

so that the transformations

$$\tilde{C} = GUG^{-1}GA + HUB = HUH^{-1}GA + HUB = HU(H^{-1}GA + B),$$
$$\tilde{C} = GUA + HUH^{-1}HB = GUA + GUG^{-1}HB = GU(A + G^{-1}HB)$$

can be made, yielding

$$UM = H^{-1}\tilde{C}H + G^{-1}\tilde{CG} = C$$

with $M = H^{-1}GAH + BH + AG + G^{-1}HBG$.  \hfill (5.4)

If now (5.2) is taken into account, then on the one hand

$$M^*H - HM = GBHG^{-1}H - HG^{-1}HBG = \mu^{-2}(GBG - GBG) = 0,$$
$$M^*G - GM = HAGH^{-1}G - GH^{-1}GAH = \mu^2(HAH - HAH) = 0$$

and on the other hand (5.4) implies

$$HCH^{-1} = \tilde{C} + HG^{-1}\tilde{CG}H^{-1} = (\mu^2/\mu^2)GH^{-1}\tilde{CG}H^{-1} + \tilde{C} = GCG^{-1}$$

or

$$H^{-1}C^*H = G^{-1}C^*G.$$  

Therefore, if (5.2) holds and if $U$ is a $(G, H)$-isometry and if there exist selfadjoint matrices $A, B$ which solve (5.3), then there exists a $(G, H)$-selfadjoint matrix $M$ such that $UM$ is a $(G, H)$-polar decomposition of $C$. In particular, it is true that $C^H = C^\circ$.  

In order to prove that the existence of a $(G, H)$-polar decomposition $UM = C$ conversely implies the existence of the matrices $A$ and $B$, assume that (5.2) holds. Then, according to Lemma 2.6, there exists a nonsingular matrix $S$ such that

$$S^*HS = J_1 \oplus J_2,$$
$$S^*GS = \mu(J_1 \ominus J_2),$$
$$S^{-1}US = U_1 \oplus U_2,$$
$$S^{-1}MS = M_1 \oplus M_2,$$
$$S^{-1}CS = C_1 \oplus C_2,$$

where $J_k$ has the form $\text{diag}(\pm 1)$ and $U_kM_k = C_k$ is a $J_k$-polar decomposition ($k = 1, 2$). Let

$$S^*\tilde{C}S^{-*} = \begin{bmatrix} \tilde{C}_{11} & \tilde{C}_{12} \\ \tilde{C}_{21} & \tilde{C}_{22} \end{bmatrix}$$

and

$$S^{-1}A^*S^{-*} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21}^{*} & A_{22} \end{bmatrix},$$
$$S^{-1}BS^{-*} = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^{*} & B_{22} \end{bmatrix}$$

be compatible partitionings. Then from (5.4) it follows that

$$U_1M_1 \oplus U_2M_2 = C_1 \oplus C_2 = 2(J_1\tilde{C}_{11}J_1 \oplus J_2\tilde{C}_{22}J_2)$$

or

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21}^{*} & A_{22} \end{bmatrix},$$
$$\begin{bmatrix} B_{11} & B_{12} \\ B_{12}^{*} & B_{22} \end{bmatrix}$$

be compatible partitionings. Then from (5.4) it follows that

$$U_1M_1 \oplus U_2M_2 = C_1 \oplus C_2 = 2(J_1\tilde{C}_{11}J_1 \oplus J_2\tilde{C}_{22}J_2)$$

or
\[ \widetilde{C}_{11} = \frac{1}{2} J_1 U_1 M_1 J_1 \quad \text{and} \quad \widetilde{C}_{22} = \frac{1}{2} J_2 U_2 M_2 J_2. \]

On the other hand, (5.3) requires \( GA + HB = U^* \widetilde{C} \) or

\[
\begin{bmatrix}
J_1 (\mu A_{11} + B_{11}) & J_1 (\mu A_{12} + B_{12}) \\
J_2 (-\mu A_{12}^* + B_{12}^*) & J_2 (-\mu A_{22} + B_{22})
\end{bmatrix}
= \begin{bmatrix}
U_1^* \widetilde{C}_{11} & U_1^* \widetilde{C}_{12} \\
U_2^* \widetilde{C}_{21} & U_2^* \widetilde{C}_{22}
\end{bmatrix},
\]

yielding the system of equations

\[
\begin{align*}
\mu A_{11} + B_{11} &= J_1 U_1^* \widetilde{C}_{11} = \frac{1}{2} M_1 J_1, \\
\mu A_{12} + B_{12} &= J_1 U_1^* \widetilde{C}_{12}, \\
-\mu A_{22} + B_{22} &= J_2 U_2^* \widetilde{C}_{22} = \frac{1}{2} M_2 J_2, \\
-\mu A_{12} + B_{12} &= \widetilde{C}_{21}^* U_2 J_2.
\end{align*}
\]

Therefore, by selecting arbitrary selfadjoint blocks \( B_{11}, B_{22} \) and setting

\[
\begin{align*}
A_{11} &= \frac{1}{\mu} \left( \frac{1}{2} M_1 J_1 - B_{11} \right) = A_{11}^*, \\
A_{12} &= \frac{1}{2\mu} (J_1 U_1^* \widetilde{C}_{12} - \widetilde{C}_{21}^* U_2 J_2), \\
A_{22} &= \frac{1}{\mu} \left( B_{22} - \frac{1}{2} M_2 J_2 \right) = A_{22}^*, \\
B_{12} &= \frac{1}{2} (J_1 U_1^* \widetilde{C}_{12} + \widetilde{C}_{21}^* U_2 J_2),
\end{align*}
\]

the two selfadjoint matrices \( A \) and \( B \) which solve (5.3) are determined. If the particular choice

\[
B_{11} = \frac{1}{4} M_1 J_1, \quad B_{22} = \frac{1}{4} M_2 J_2
\]

is made, then

\[
A_{11} = \frac{1}{4\mu} M_1 J_1, \quad A_{22} = -\frac{1}{4\mu} M_2 J_2
\]

and thus

\[
\begin{align*}
A &= \frac{1}{4} (G^{-1} U^* \widetilde{C} + \widetilde{C}^* U G^{-1}) - \frac{1}{4} M G^{-1}, \\
B &= \frac{1}{4} (H^{-1} U^* \widetilde{C} + \widetilde{C}^* U H^{-1}) - \frac{1}{4} M H^{-1},
\end{align*}
\]

which follows from (5.5). Summarising, the following result is proved.

**Lemma 5.1** (\( F = \mathbb{R} \) or \( F = \mathbb{C} \)). Let \( G, H \in \mathbb{F}^{n \times n} \) be nonsingular selfadjoint matrices which satisfy (5.2). Moreover, let \( U \in \mathbb{F}^{n \times n} \) be a \( (G, H) \)-isometry and let \( \widetilde{C} \in \mathbb{F}^{n \times n} \). Then the following statements are equivalent:

(i) There exist selfadjoint matrices \( A, B \in \mathbb{F}^{n \times n} \) such that

\[
G A + H B = \widetilde{C}.
\]

(ii) There exists a \( (G, H) \)-selfadjoint matrix \( M \in \mathbb{F}^{n \times n} \) such that

\[
U M = G^{-1} \widetilde{C} G + H^{-1} \widetilde{C} H.
\]
Using this lemma, the necessary condition (5.3) for solving the Procrustes problem (5.1) under the assumption (5.2) finally becomes

\[
C = UM \quad \text{with} \quad C = YX^*H + G^{-1}HYX^*G \quad \text{and} \quad C^H = C^G, \quad U^H = U^G = U^{-1}, \quad M^H = M^G = M. \tag{5.6}
\]

Thus the solution of the problem can be determined by a \((G,H)\)-polar decomposition of the matrix \(C\).

Again, it remains to determine the particular \((G,H)\)-polar decomposition (if existent) which leads to the optimum congruence. For this, let \(UM\) be a \((G,H)\)-polar decomposition of the matrix \(C\). Moreover, let \(S\) be a nonsingular matrix such that (5.5) holds and let

\[
S^{-1}X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \quad \text{and} \quad S^{-1}Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}
\]

be compatible partitionings. Then on the one hand from (5.3), (5.5a), (5.5b) it follows that

\[
\begin{bmatrix} \tilde{C}_{11} & \tilde{C}_{12} \\ \tilde{C}_{21} & \tilde{C}_{22} \end{bmatrix} = S^{-2}CS^{-*} = (S^*HS)(S^{-1}Y)(S^{-1}X)^* = \begin{bmatrix} J_1Y_1X_1^* & J_1Y_2X_2^* \\ J_2Y_1X_2^* & J_2Y_2X_2^* \end{bmatrix}.
\]

so that according to (5.5d)

\[
U_kM_k = C_k = 2J_k\tilde{C}_{kk}J_k = 2Y_kX_k^*J_k \quad \text{for} \quad k = 1, 2.
\]

On the other hand we find from the initial equation (5.1b)

\[
\begin{align*}
f(U) &= \text{tr} \left( \left[ (S^{-1}US)(S^{-1}X) - (S^{-1}Y) \right]^* (S^*HS) \left[ (S^{-1}US)(S^{-1}X) - (S^{-1}Y) \right] \right) \\
&= \text{tr} \left( \left[ U_1X_1 - Y_1 \right]^* (J_1 \oplus J_2) \left[ U_1X_1 - Y_1 \right] \right) \\
&= \sum_k \text{tr} \left[ (U_kX_k - Y_k)^*J_k(U_kX_k - Y_k) \right].
\end{align*}
\]

Now, using the canonical forms of the pairs \((C_k^j, C_k, J_k) = (M_k^j, J_k)\) and \((M_k, J_k)\) the argumentation from Section 4 can be applied twice, showing that the optimum congruence is achieved when \(U_kM_k\) are semidefinite \(J_k\)-polar decompositions of the matrices \(C_k\). If in this case we set

\[
\begin{bmatrix} X_1' \\ X_2' \end{bmatrix} = \begin{bmatrix} U_1X_1 \\ U_2X_2 \end{bmatrix} = S^{-1}UX \quad \text{and} \quad \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = S^{-1}Y, \tag{5.7a}
\]

then the orthogonal or unitary Procrustes problems

\[
\varphi_k(T_k) = \text{tr}[(T_kX_k' - Y_k)^*(T_kX_k' - Y_k)] \to \min \quad \text{with} \quad T_k^*T_k = I, \tag{5.7b}
\]
are solved for $T_k = I$. Moreover, if $H > 0 (H < 0)$, then $J_k = I (J_k = -I)$, so that a solution then always exists (see Example 2.8). Summarising, the result can be expressed by the following theorem.

**Theorem 5.2** (Solution of the $(G, H)$-isometric Procrustes problem). A solution of the $(G, H)$-orthogonal or $(G, H)$-unitary Procrustes problem (5.1) under the assumption (5.2) exists if and only if the matrix $C = YX^*H + G^{-1}HYX^*G$ admits an $H$-semidefinite $(G, H)$-polar decomposition. In this case the $(G, H)$-isometry $U$ contained in such a decomposition $C = UM$ optimises the function $f$. Moreover, $X' = UX$ and $Y$ are at optimum congruence in the sense, that the orthogonal or unitary Procrustes problems (5.7) are solved for $T_k = I$.

6. Numerical computation of $H$- and $(G, H)$-polar decompositions

This final section provides a method for the numerical computation of $H$- and $(G, H)$-polar decompositions. The method is obtained via generalisation of the Newton iteration for computing ordinary polar decompositions which is closely related to the matrix sign function [9,11]. An extension for the computation of the semidefinite factorisations will also be described.

Let $A \in \mathbb{F}^{n \times n}$ be nonsingular. Then the iteration method

$$X_{k+1} = \frac{1}{2} (X_k + X_k^{-*}), \quad X_0 = A, \quad k = 0, 1, \ldots$$

quadratically converges to the isometric factor of an ordinary polar decomposition $A = UM$ with $U^* = U^{-1}$ and $M^* = M$. The optimum selfadjoint factor corresponding to the approximate isometry

$$\tilde{U} = X_k \quad \text{with} \quad \|X_k^*X_k - I\| < \varepsilon \quad \text{(Frobenius norm)}$$

is given by [12, Lemma 2.1]

$$\tilde{M} = \frac{1}{2} (A^*\tilde{U} + \tilde{U}^*A),$$

and, furthermore, the particular polar decomposition is computed for which $M$ is positive definite. By substituting the adjoints in (6.1) with $H$-adjoints we obtain the iteration method

$$X_{k+1} = \frac{1}{2} (X_k + X_k^{-H}), \quad X_0 = A, \quad k = 0, 1, \ldots, \quad \tilde{U} = X_k \quad \text{with} \quad \|X_k^H X_k - I\| < \varepsilon \quad \text{and} \quad \tilde{M} = \frac{1}{2} (A^H\tilde{U} + \tilde{U}^H A),$$

for which it will now be shown that it serves for the computation of an $H$-polar decomposition.
Let $A$ be a matrix for which $\sigma(AH^HA) \subseteq \mathbb{C}\setminus(-\infty, 0]$. Then, according to Proposition 2.1, $A$ admits an $H$-polar decomposition $A = U_0M_0$ such that $\sigma(M_0)$ lies in the open right complex half-plane. Now applying (6.2) to $A$ we obtain
\[
\begin{align*}
2X_1 &= (U_0M_0 + (U_0M_0)^{-H}) = (U_0M_0 + (M_0^H U_0^H)^{-1}) \\
&= (U_0M_0 + (M_0U_0^{-1})^{-1}) = (U_0M_0 + U_0M_0^{-1}) = U_0(M_0 + M_0^{-1})
\end{align*}
\]
or
\[
X_1 = U_0M_1 \text{ with } M_1 = \frac{1}{2}(M_0 + M_0^{-1}) = M_0^H,
\]
from which it follows that
\[
X_{k+1} = U_0M_{k+1} \text{ with } M_{k+1} = \frac{1}{2}(M_k + M_k^{-1}) = M_k^H.
\]
Moreover, $\text{Re} \lambda > 0$ for all $\lambda \in \sigma(M_0)$ according to [16] implies
\[
\lim_{k \to \infty} M_k = I,
\]
so that finally we have
\[
\lim_{k \to \infty} X_k = U_0.
\]

**Theorem 6.1.** Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and let $A \in F^{n \times n}$ be a matrix for which $\sigma(AH^HA) \subseteq \mathbb{C}\setminus(-\infty, 0]$. Then the algorithm (6.2) applied to $A$ computes the particular $H$-polar decomposition $A = UM$, for which $\sigma(M)$ lies in the open right complex half-plane.\(^5\)

If now $H$-isometric Procrustes problems are to be solved, then this requires semi-definite ($HM \succeq 0$) or definite ($HM > 0$) $H$-polar decompositions, respectively. However, an $H$-polar decomposition $A = UM$ computed with algorithm (6.2) has this property only in the case $H > 0$. In the case $H < 0$ the decomposition $A = (-U)(-M)$ can still be used, but in any other case $HM$ is indefinite. Therefore, the problem is to find a modification or an extension of algorithm (6.2) with the help of which the computation of definite polar decompositions is possible, too. For this, the following theorem on a simplified canonical form is needed which could simply be deduced from [6, Theorem I.3.3]. We will provide another proof which helps to solve the present problem.

**Theorem 6.2 (Simplified canonical form).** Let $H \in \mathbb{C}^{n \times n}$ be Hermitian and nonsingular and let $A \in \mathbb{C}^{n \times n}$ be $H$-Hermitian and diagonalisable. Then there exists a nonsingular matrix $S \in \mathbb{C}^{n \times n}$ such that

\[
S^{-1}AS = \bigoplus_{j=1}^{r} \begin{bmatrix} \lambda_j I_{p_j} \\ \lambda_j I_{q_j} \end{bmatrix} \oplus \bigoplus_{j=r+1}^{s} \begin{bmatrix} \lambda_j I_{p_j} \\ \overline{\lambda_j} I_{p_j} \end{bmatrix}.
\]

\(^5\) A similar result has been derived independently by Higham [10, Theorem 5.2].
S^*HS = \bigoplus_{j=1}^{r} \begin{bmatrix} I_{p_j} & -I_{q_j} \end{bmatrix} \oplus \bigoplus_{j=r+1}^{s} \begin{bmatrix} I_{p_j} & I_{q_j} \end{bmatrix},

where \( \lambda_j \in \mathbb{R} \) for \( 1 \leq j \leq r \) and \( \lambda_j \in \mathbb{C} \setminus \mathbb{R} \) for \( r + 1 \leq j \leq s \).

**Proof.** Let \( R \) be a nonsingular matrix consisting of eigenvectors of \( A \). Then it is easily derived from [6, Theorem I.2.5], that

\[
R^{-1}AR = \bigoplus_{j=1}^{r} A_j \oplus \bigoplus_{j=r+1}^{s} \begin{bmatrix} A_j & \bar{A_j} \end{bmatrix},
\]

\[
R^*HR = \bigoplus_{j=1}^{r} H_j \oplus \bigoplus_{j=r+1}^{s} \begin{bmatrix} H_j & H_j^* \end{bmatrix},
\]

where \( A_j \) contain the eigenvalues, \( H_j \) are blocks of compatible sizes, and a suitable sorting of the eigenvectors is assumed. Let

\[
H_j = P_j \Omega_j P_j^* = H_j^* \quad \text{for } 1 \leq j \leq r \quad \text{and} \quad H_j = P_j \Omega_j Q_j^* \quad \text{for } r + 1 \leq j \leq s
\]

be eigenvalue or singular value decompositions, respectively. Then \( \Omega_j \) are nonsingular real diagonal matrices and for \( r + 1 \leq j \leq s \) their diagonal elements are positive. Moreover, \( P_j \) and \( Q_j \) are unitary. Hence, for

\[
S = R \left( \bigoplus_{j=1}^{r} P_j [\Omega_j]^{-1/2} \oplus \bigoplus_{j=r+1}^{s} \begin{bmatrix} P_j \Omega_j^{-1/2} & 0 \\ 0 & Q_j \Omega_j^{-1/2} \end{bmatrix} \right)
\]

the statement of the theorem holds, when again a suitable sorting is made. (The operations on \( \Omega_j \) are to be applied to its diagonal elements.) \( \square \)

Now, let \( A \in \mathbb{F}^{n \times n} \) (be nonsingular and) admit a definite \( H \)-polar decomposition. Then, according to Proposition 2.2, \( A^H \mathcal{A} \) must be diagonalisable and \( \sigma(A^H \mathcal{A}) \subset (0, \infty) \). Thus an \( H \)-polar decomposition \( A = UM \) can be obtained with algorithm (6.2) where \( M \) is also diagonalisable and has only positive eigenvalues. Moreover, Theorem 6.2 gives the possibility to compute a nonsingular matrix \( S \in \mathbb{F}^{n \times n} \) such that

\[
S^{-1}MS = J = \bigoplus_{j=1}^{r} \begin{bmatrix} \lambda_j I_{p_j} & -I_{q_j} \end{bmatrix} \quad \text{and} \quad S^*HS = Z = \bigoplus_{j=1}^{r} \begin{bmatrix} I_{p_j} & -I_{q_j} \end{bmatrix},
\]

(6.3)

where \( \lambda_j > 0 \) for \( 1 \leq j \leq r \) are the eigenvalues of \( M \). For the matrices defined by

\[
V = SS^{-1} \quad \text{and} \quad K = SZS^{-1}
\]

(6.4a)
it is obviously true that $M = VK$ and, using $S^{-1} = ZS^\ast H$, it can be verified that $V$ is an $H$-isometry and that $K$ is $H$-selfadjoint. Furthermore, $HK = (S^{-\ast} ZS^{-1}) \times (SZJS^{-1}) = S^{-\ast} JS^{-1}$ is positive definite, so that $M = VK$ and thus also

$$A = (UV)K$$

(6.4b)

are definite $H$-polar decompositions.

**Algorithm 6.3.** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let $A \in \mathbb{F}^{n \times n}$ admit a definite $H$-polar decomposition. Then this decomposition can be computed with the following steps:

1. Compute the $H$-polar decomposition $A = UM$ using (6.2).
2. Compute the simplified canonical form of the pair $(M, H)$ using (6.3).
3. Compute the definite $H$-polar decomposition $A = (UV)K$ using (6.4).

The algorithms also apply for computing $(G, H)$-polar decompositions. Indeed, if $G$, $H$ and $A$ are as in Lemma 2.7, then every $H$-polar decomposition with $\sigma(M) \subset \mathbb{C}^+$ is also a $G$-polar decomposition and vice versa.6 Thus algorithm (6.2) applied to $A$ converges to a $(G, H)$-polar decomposition. If $H$ is positive or negative definite, then this decomposition $A = UM$ or $A = (-U)(-M)$, respectively, is also $H$-definite. However, if $H$ is indefinite, then an $H$-semidefinite $(G, H)$-polar decomposition must be obtained using Algorithm 6.3 applied to the blocks $A_1$ and $A_2$ which exist according to Lemma 2.6. In fact, Example 2.10 shows, that in this case a definite $H$-polar decomposition of $A$ may not be a $G$-polar decomposition, too. In addition to this, the following remarks are in order:

(a) Algorithm (6.2) represents the basic form of the Newton iteration. It may be improved with factors for convergence acceleration analogously to [12, Section 2].

(b) The $H$-orthogonalisation of the eigenspaces in Theorem 6.2 can also be performed with generalisations of the Gram–Schmidt orthogonalisation method as is described in [13, Method 3.24]. A sorting algorithm for the grouping of the eigenvalues is also given there.

(c) A further algorithm for the computation of $H$-polar decompositions of matrices $A$ for which $A^H A$ is diagonalisable is given in [13, Method 3.25]. This method is based on the computation of the simplified canonical form of the pair $(A^H A, H)$ and is therefore not always stable if $A$ is ill-conditioned. However, it has the advantage, that it also applies to singular matrices.

In order to be able to assess the numerical properties of the methods, a corresponding implementation was tested using the programming language C and double precision floating point numbers. For this purpose the canonical forms

---

6 This statement also holds when $\rho H - G$ is a non-defective Hermitian pencil which can be shown using a corresponding generalisation of Lemma 2.7.
\[ K = \bigoplus J_p(\lambda) \oplus J_q(\alpha) \oplus J_r(\beta), \quad Z = \bigoplus \varepsilon_\alpha Z_q \oplus \varepsilon_\beta Z_r, \]

with \( \lambda \in \mathbb{C} \setminus \mathbb{R}, \alpha, \beta \in \mathbb{R} \setminus \{0\} \), \( \varepsilon_\alpha = +1, \varepsilon_\beta = -1, p = q = r = 10, \text{i.e. } n = 40 \), were specified to test algorithm (6.2), and the canonical forms

\[ K = \bigoplus_{j=1}^4 \lambda_j I_{p_j}, \quad Z = \text{diag}_n(\pm 1) \]

with \( \lambda_j \in \mathbb{R} \setminus \{0\}, p_j = 10, \text{i.e. } n = 40 \), were specified to test Algorithm 6.3.

Using randomly chosen eigenvalues, transformations \( S \in \mathbb{C}^{n \times n} \) and \( Z \)-isometries \( T \in \mathbb{C}^{n \times n} \), test examples of the kind

\[ A = S^{-1}TKS, \quad H = S^*ZS \in \mathbb{C}^{n \times n} \]

were constructed, always based on normally distributed random numbers from the interval \([-2, 2]\). The magnitudes of the eigenvalues of \( K \) were at least 0.2. Finally, \( H \)-polar decompositions \( UM \) or definite \( H \)-polar decompositions \( UM' = (UV)K \) of the test matrices \( A \) were computed, whose numerical accuracy is estimated via the residuals

\[ r_A = \| A - UM \|, \quad r_M = \| M^*H - HM \|, \quad r_U = \| U^*HU - H \| \]

and the condition number \( c_U = \| U \| \| U^{-1} \| \), respectively.

We used the Gauss–Jordan method [18, Kapitel 4.2] for matrix inversion in algorithm (6.2). The eigenvalues \( \lambda_j \) and eigenvectors \( R \) used in the proof of Theorem 6.2 were computed with the QR algorithm [8, Chapter 7.5] and the Hermitian eigenvalue problem \( H_j = P_j \Omega_j P_j^* \), also contained in this proof, was solved with the symmetric QR algorithm [8, Chapter 8.3].

The results of two statistical experiments with 50 repetitions are shown in Tables 1 and 2. There \( its \) is the number of iterations, \( \mu \) the (empirical) mean value, \( \sigma^2 \) the (empirical) variance and \( \min/\max \) specify the respective minimum and maximum value. The machine accuracy and the tolerance parameter \( \varepsilon \) from (6.2) were given as

\[ \varepsilon_{\text{mach}} \approx 2.22 \cdot 10^{-16} \quad \text{and} \quad \varepsilon = 10^{-8}. \]

The tables show that the iteration in most cases required only between 6 and 8 steps. Taking into account that the residuals are absolute errors of 40 \( \times \) 40 matrices, both algorithms seem to be appropriate for computing the respective \( H \)-polar decompositions.

### Table 1
Results of a statistical experiment with algorithm (6.2)

<table>
<thead>
<tr>
<th>( m = 50 )</th>
<th>( \mu )</th>
<th>( \sigma^2 )</th>
<th>( \min )</th>
<th>( \max )</th>
<th>( 10^\mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( its )</td>
<td>7.720</td>
<td>1.798</td>
<td>6</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td>( \log r_A )</td>
<td>-8.781</td>
<td>3.201</td>
<td>-11.434</td>
<td>-2.825</td>
<td>1.659e-09</td>
</tr>
<tr>
<td>( \log r_M )</td>
<td>-8.667</td>
<td>1.060</td>
<td>-10.367</td>
<td>-6.249</td>
<td>2.152e-09</td>
</tr>
<tr>
<td>( \log r_U )</td>
<td>-8.621</td>
<td>1.451</td>
<td>-10.664</td>
<td>-4.760</td>
<td>2.394e-09</td>
</tr>
<tr>
<td>( \log c_U )</td>
<td>4.812</td>
<td>1.523</td>
<td>3.203</td>
<td>8.874</td>
<td>64.916</td>
</tr>
</tbody>
</table>
Table 2
Results of a statistical experiment with algorithm (6.3)

<table>
<thead>
<tr>
<th></th>
<th>μ</th>
<th>σ²</th>
<th>min</th>
<th>max</th>
<th>10μ⁴</th>
</tr>
</thead>
<tbody>
<tr>
<td>m = 50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>its</td>
<td>6.313</td>
<td>0.262</td>
<td>5</td>
<td>7</td>
<td>–</td>
</tr>
<tr>
<td>log rₐ’</td>
<td>-7.693</td>
<td>1.131</td>
<td>-9.275</td>
<td>-4.667</td>
<td>2.029e-08</td>
</tr>
<tr>
<td>log rᵤ’</td>
<td>-7.492</td>
<td>1.002</td>
<td>-9.060</td>
<td>-4.424</td>
<td>3.222e-08</td>
</tr>
<tr>
<td>log cᵤ’</td>
<td>4.637</td>
<td>0.568</td>
<td>3.505</td>
<td>6.804</td>
<td>43.316</td>
</tr>
</tbody>
</table>

7. Conclusions

Theorem 3.1 generalises a method for constructing vectors from given values of a quadratic form, and Theorem 4.1 or Theorem 5.2, respectively, give solutions for the Procrustes problems (4.1) and (5.1). These results constitute a foundation for multidimensional scaling in an environment of indefinite scalar products. In preparation of these investigations the concept of (G,H)-polar decompositions is introduced which is particularly useful for solving the problem (5.1).

Moreover, the algorithm (6.2) for the numerical computation of H- or (G,H)-polar decompositions is given. This method represents a basic approach which could be refined in further research. With the extension to Algorithm 6.3 the practical solution of indefinite Procrustes problems is possible, too.

Acknowledgments

I would like to thank Peter Benner, Technische Universität Chemnitz, for reading the manuscript and for his suggestions which enhanced the quality of my presentation. Moreover, I thank Leiba Rodman, College of William and Mary, and an anonymous referee for their valuable hints.

References