L^2 Approximation by Euclidean motions *

Johannes Wallner

Abstract

For a Euclidean space V, we consider L^2 approximation of mappings $V \to V$ by Euclidean congruence transformations. It turns out that the problem essentially reduces to approximating linear mappings by orthogonal linear mappings in the L^2 sense, and that the latter problem is solved by a certain polar decomposition. The medial axes of the orthogonal and special orthogonal groups within the linear space $\mathrm{End}(V)$ are computed.

Introduction

It is the aim of this paper to study the best approximation of 'deformations' by Euclidean congruence transformations. When studying the 'difference' of mappings it is natural to employ L^2 spaces, so this approximation is by considering the action of a deformation on a certain distribution of mass (i.e., a positive Borel measure in a linear space V) and looking for the Euclidean congruence transformation which is closest to the given deformation in the sense of $L^2(V, V)$.

The restriction of the L^2 metric to the orthogonal linear group has been considered in mechanics, where it provides a left invariant Riemannian metric on the Lie group of motions of a rigid body (identified with the measure referred to above); the geodesics of the metric coincide with the trajectories of the free motion of this rigid body. References to these aspects of the L^2 metric are given in [3]. This paper studies the integrability of the equation of geodesics.

In Euclidean three-space the problem of approximating a given deformation of a finite point cloud by a congruent and positively oriented copy of that cloud has been considered in [2]. The motivation lies in registering a Euclidean displacement to a 'deformation' very close to a Euclidean

*Technical Report No. 93, Institut für Geometrie, TU Wien

congruence transformation, and which stems from imprecisely measured positions of data points. The paper [2] uses quaternions and arrives at a certain eigenvalue problem, which (unsurprisingly) will also be the case in this paper.

The connections of the best approximation with the polar decomposition as described in Prop. 6, and the computation of the medial axes of $\mathrm{O}(V)$ and $\mathrm{SO}(V)$ within $\mathrm{End}(V)$ (Prop. 8 and Prop. 9) are perhaps new even in the three-dimensional case.

Our motivation for considering the problem of approximating 'deformation' by Euclidean congruence transformations comes from a motion design problem of Computer-Aided Geometric Design (cf. [1]): The selection of a finite number of points in a rigid body and designing a path for each of them does of course not lead to a Euclidean motion unless these paths respect the Euclidean distance between the points. The collection of these paths will however constitute a smooth path in $L^2(V,V)$, and projecting this path onto the orthogonal group will yield a smooth Euclidean motion. A particular and important instance where path generation is important is the variational design of motions constrained by surface-surface contact, such as appears in the mathematical treatment of 5-axis NC machining.

It is therefore important to know where the orthogonal projection onto the motion group will be smooth or even analytic. The answer is given in two steps: First the projection onto the space of affine transformations is smooth, and the projection from this intermediate space onto the group of Euclidean congruence transformations is smooth precisely in $\mathrm{GL}(V)$ (cf. p. 10).

This paper is organized as follows: In the first section we consider approximation of linear mappings by orthogonal linear mappings; the second section contains the general approximation problem. Finally a short third section enumerates some useful formulae for the case of point clouds.

1 L^2 distance of linear mappings

We consider an n-dimensional real linear space V, and the linear space $\mathrm{End}(V)$ of its linear endomorphisms. V shall be endowed with a positive definite scalar product $\langle\ ,\ \rangle$ and the corresponding norm $\|\cdot\|$. Further a positive Borel measure μ is given in V. The Banach space $L^2(V,V)$ consists of all measurable mappings $F : V \to V$ with

$$\|F\|^2 = \int \|Fx\|^2 \, d\mu(x) < \infty, \tag{1}$$

with the usual identification of mappings which are equal almost everywhere. If $(e_1, \dots, e_n)$ is an orthonormal basis, we define, for $F \in L^2(V,V)$ the functions $F_j : V \to \mathbb{R}$, $F_j(x) = \langle F(x), e_j \rangle$. From

$$\|F\|^2_{L^2(V,V)} = \int \sum |F_j(x)|^2 = \sum \|F_j\|^2_{L^2(V,\mathbb{R})} \tag{2}$$

follows that $L^2(V,V)$ is isometric to the orthogonal direct sum of $\dim(V)$ copies of $L^2(V,\mathbb{R})$.

In order to introduce a scalar product in $\mathrm{End}(V)$, we assume that

$$\mathrm{End}(V) \subset L^2(V,V) \tag{3}$$

and use the restriction of the L^2 scalar product to $\mathrm{End}(V)$. This means that

$$\|A\|^2 = \int \|Ax\|^2 \, d\mu(x), \quad \langle A, B \rangle = \int \langle Ax, Bx \rangle \, d\mu(x). \tag{4}$$

An important special case is that μ is a finite sum of point masses:

$$\mu = \sum_{x \in X} \delta_x \Longrightarrow \|A\|^2 = \sum_{x \in X} \|Ax\|^2, \quad \langle A, B \rangle = \sum_{x \in X} \langle Ax, Bx \rangle. \tag{5}$$

The distance of linear mappings is given by the obvious definition

$$d(A,B) = \|A - B\|. \tag{6}$$

The adjoint of a linear mapping A is denoted by A^*. The scalar product in $\mathrm{End}(V)$ has the following elementary properties:

$$\langle AB, C \rangle \;=\; \langle B, A^* C \rangle, \quad \langle A, B \rangle = \langle I, A^* B \rangle, \tag{7}$$

which follow directly from the defining equation $\langle x, Ay \rangle = \langle A^* x, y \rangle$ of A^*. Consider the *orthogonal group*

$$\mathrm{O}(V) = \{ A \in \mathrm{End}(V) \mid A^{-1} = A^* \} \tag{8}$$

It is contained in a level set of $\| \cdot \|$, because

$$P \in \mathrm{O}(V) \implies \langle P, P \rangle = \langle I, P^* P \rangle = \langle I, I \rangle. \tag{9}$$

The inertia tensor

The inertia tensor of μ is the unique symmetric and nonnegative $J \in \mathrm{End}(V)$ which fulfills

$$\int \langle x,v\rangle \langle x,w\rangle \, d\mu(x) = \langle v, Jw\rangle. \tag{10}$$

The coordinate matrix J_{ij} of J with respect to an orthonormal basis $(e_1,\ldots,e_n)$ has the entries

$$J_{ij} = \langle e_i, Je_j\rangle = \int \langle x,e_i\rangle\langle x,e_j\rangle \, d\mu(x) = \int x_i x_j \, d\mu(x). \tag{11}$$

Here x_i denotes the i-th coordinate of x with respect to the basis $(e_1,\ldots,e_n)$. If $e_1,\ldots,e_n$ are eigenvectors of J, then the matrix $(J_{ij})_{i,j=1}^n$ is diagonal, and the sums $J_{ii}+J_{jj}$ are called the principal moments of inertia of the mass distribution μ. It is easy to verify that without changing the norm $\|\cdot\|$ in $\mathrm{End}(V)$, we may replace μ by the sum

$$\mu' = \delta_{\sqrt{J_{11}}e_1} + \cdots + \delta_{\sqrt{J_{nn}}e_n} \tag{12}$$

of point masses of unit weight at the locations $\sqrt{J_{ii}}e_i$, where $e_1,\ldots,e_n$ is an orthonormal basis of eigenvectors of J, and $J_{ii} = q(e_i,e_i)$.

Proposition 1 *If J is nonsingular, the L^2 scalar product of linear mappings has the properties*

$$\langle I, MJ^{-1}\rangle = \mathrm{tr}M, \quad \langle A,B\rangle = \mathrm{tr}(A^*BJ) = \mathrm{tr}(B^*AJ) \tag{13}$$

Proof: In view of Equ. (7) it is sufficient to show $\langle I, MJ^{-1}\rangle = \mathrm{tr}M$. The decomposition $M = \frac{1}{2}(M+M^*) + \frac{1}{2}(M-M^*)$, shows that it is sufficient to treat the cases of symmetric and skew-symmetric M separately. As to symmetric linear mappings, they are linear combinations of orthogonal projections $x \mapsto \langle x,v\rangle v$ with $\langle v,v\rangle = 1$, so it is sufficient to show Equ. (13) for $M = \langle v,\cdot\rangle v$:

$$\begin{aligned}\langle I, MJ^{-1}\rangle &= \int \langle x, \langle v, J^{-1}x\rangle v\rangle \, d\mu(x) = \int \langle v,x\rangle\langle J^{-1}v,x\rangle d\mu(x) \\ &= \langle J^{-1}v, Jv\rangle = \langle v,v\rangle = 1 = \mathrm{tr}M.\end{aligned} \tag{14}$$

As to skew-symmetric endomorphisms, they are linear combinations of mappings of the form $Mx = \langle a,x\rangle b - \langle b,x\rangle a$:

$$\begin{aligned}\langle I, MJ^{-1}\rangle &= \int \left\langle x, \langle a, J^{-1}x\rangle b - \langle b, J^{-1}x\rangle a \right\rangle d\mu(x) \\ &= \int (\langle J^{-1}a,x\rangle\langle x,b\rangle - \langle J^{-1}b,x\rangle\langle x,a\rangle)\, d\mu(x) \\ &= \langle J^{-1}a, Jb\rangle - \langle J^{-1}b, Ja\rangle = 0 = \mathrm{tr}M.\end{aligned} \tag{15}$$

□

Proposition 2 *J is nonsingular if and only if μ is not concentrated in a subspace U of dimension $< n$. For any such subspace, $U^\perp \leq \ker J$.*

Proof: Assume that μ is concentrated in U and $v \in U^\perp$. Then $\int_V f(x)\, d\mu(x) = \int_U f(x)\, d\mu(x) = 0$, because $\langle v, x\rangle = 0$ for $x \in U$. By Equ. (10), $v \in \ker J$.

Conversely we assume that $v \in \ker J \setminus 0$, which implies that $\int_V \langle v, x\rangle^2\, d\mu(x) = 0$. At distance $\geq c$ from $v^\perp$ the function $|\langle v, x\rangle|$ is bounded from below by $c\|v\|$. Thus no measurable subset of $V \setminus v^\perp$ has positive measure, and the support of μ is contained in $v^\perp$. □

We will always assume that J is nonsingular.

Left and right invariance

The L^2 norm in $\mathrm{End}(V)$ is left $\mathrm{O}(V)$-invariant in the sense that for all $P \in O(V)$ we have

$$\langle PA, PB\rangle = \langle A, P^* PB\rangle = \langle A, B\rangle. \tag{16}$$

Further, there is a subgroup O^* of $\mathrm{GL}(V)$ with respect to which the norm is right invariant in the sense that

$$P \in \mathrm{O}^* \iff \langle AP, BP\rangle = \langle A, B\rangle \quad \text{for all } A, B. \tag{17}$$

Proposition 3 *The group O^* is given by*

$$P \in \mathrm{O}^* \iff PJP^* = J \tag{18}$$

Proof: The condition of right invariance reads

$$\mathrm{tr}(A^*BJ) = \langle A, B\rangle = \langle AP, BP\rangle = \mathrm{tr}(P^*A^*BPJ) = \mathrm{tr}(A^*BPJP^*). \tag{19}$$

for all A, B. □

Subspaces orthogonal to $\mathrm{O}(V)$.

As $\mathrm{O}(V)$ is a closed subset of $\mathrm{End}(V)$, for all $A \in \mathrm{End}(V)$ there is a $P \in \mathrm{O}(V)$ such that the distance $d(A, P)$ is minimal. In this case $A \in T_P\mathrm{O}(V)^\perp$. Conversely, if $A \in T_P\mathrm{O}(V)^\perp$, then P makes the distance $d(A, P)$ stationary. Recall that the tangent space of $\mathrm{O}(V)$ at $P \in \mathrm{O}(V)$ is given by

$$A \in T_P\mathrm{O}(V) \iff P^{-1}A \text{ is skew-symmetric.} \tag{20}$$

Proposition 4 *The affine subspace orthogonal to* $\mathrm{O}(V)$*'s tangent space at* $P \in \mathrm{O}(V)$ *is given by*

$$A \in T_P\mathrm{O}(V)^{\perp} \iff (P^{-1}AJ) \text{ is symmetric.} \tag{21}$$

Proof: The L^2 scalar product in $\mathrm{End}(V)$ is left $\mathrm{O}(V)$-invariant, so it is sufficient to show Equ. (21) for the case $P = I$. As $\dim \mathrm{O}(V) = n(n-1)/2$, we have $\dim T_I\mathrm{O}(V)^{\perp} = n(n+1)/2$. The linear subspace $\{A \mid (AJ) = (AJ)^*\}$ of $\mathrm{End}(V)$ has the same dimension. It is therefore enough to show "$\Longrightarrow$" in Equ. (21). So we assume that $JA^* = AJ$ and verify $\langle A, X\rangle = 0$ for all all X with $X + X^* = 0$:

$$\begin{aligned}\langle A, X\rangle &= \mathrm{tr}(X^*AJ) = \mathrm{tr}(JA^*X^*) = -\,\mathrm{tr}(JA^*X) \\ &= -\,\mathrm{tr}(XJA^*) = -\,\mathrm{tr}(AJX^*) = -\,\mathrm{tr}(X^*AJ) = -\langle A, X\rangle\end{aligned} \tag{22}$$

The preceding equation shows that the (linear) orthogonal complement of the tangent space at I has the equation $AJ = (AJ)^*$. As this equation is fulfilled also by $-I$, the affine orthogonal space contains $0 \in \mathrm{End}(V)$ and is in fact a linear subspace of $\mathrm{End}(V)$. □

In the following, if A is symmetric and nonnegative, the symbol $\sqrt{A}$ denotes any symmetric linear endomorphism whose square equals A. The symbol $\sqrt[+]{A}$ denotes the unique nonnegative square root.

Proposition 5 *The subset of all* $P \in \mathrm{O}(V)$ *where* $d(\cdot, A)$ *is stationary (i.e,* $A \in T_P\mathrm{O}(V)^{\perp}$*) is given by*

$$A \in T_P\mathrm{O}(V)^{\perp} \iff AJ = P\sqrt{JA^*AJ} \tag{23}$$

For nonsingular A*, this is equivalent to*

$$P \in AJ(\sqrt{JA^*AJ})^{-1} \tag{24}$$

Proof: By Prop. 4, the left hand side of Equ. (23) is equivalent to $A = PB$ with BJ symmetric. This implies that

$$(BJ)^2 = (BJ)^*(BJ) = (P^{-1}AJ)^*(P^{-1}AJ) = JA^*PP^*AJ = JA^*AJ. \tag{25}$$

If the right hand side of Equ. (23) is fulfilled, then clearly the same computation as in Equ. (25) shows that $(BJ)^*(BJ) = JA^*AJ$, i.e., BJ is symmetric and therefore $B \in T_I\mathrm{O}(V)^{\perp}$. As $A = PB$, we have $A \in T_P\mathrm{O}(V)^{\perp}$.

In the case of nonsingular A, also JA^*AJ and its square roots are nonsingular. We show that P, which is then uniquely determined by A and B, is contained in $\mathrm{O}(V)$:

$$\begin{aligned} P^*P &= (B^{-1})^*A^*AB^{-1} = (B^{-1})^*J^{-1}JA^*AJJ^{-1}B^{-1} \\ &= ((BJ)^{-1})^*(JA^*AJ)(BJ)^{-1} = (BJ)^{-1}(BJ)^2(BJ)^{-1} = I. \end{aligned} \tag{26}$$

□

Note that the computation of Prop. 5 is nothing but the polar decomposition of AJ, with the difference that the usual polar decomposition requires the positive square root.

Minimizing the distance to $\mathrm{O}(V)$

In the following, the set of eigenvalues of a linear mapping A will be denoted by the symbol $\mathrm{spec}(A)$. If A is normal, P_λ shall denote the orthogonal projection onto the eigenspace of the eigenvalue λ, so that $A = \sum_{\lambda\in\mathrm{spec}\,A} \lambda P_\lambda$. If λ is no eigenvalue, we let $P_\lambda = 0$. The preceding equation could also be written in the form $A = \sum_{\lambda\in\mathbb{R}} \lambda P_\lambda$.

Proposition 6 *The element P of $\mathrm{O}(V)$ closest to $A \in \mathrm{End}(V)$ is unique if A is nonsingular, and not unique otherwise. In the nonsingular case it is given by*

$$P = AJ(\sqrt[+]{JA^*AJ})^{-1} \tag{27}$$

Proof: By left invariance, it is sufficient to consider the case that A is contained in $T_I\mathrm{O}(V)^\perp$, i.e., the case of AJ symmetric. We use the spectral decomposition of AJ and compute

$$AJ = \sum_{\lambda>0} \lambda(P_\lambda - P_{-\lambda}), \quad JA^*AJ = \sum_{\lambda>0} \lambda^2(P_\lambda + P_{-\lambda}) \tag{28}$$

$$\Longrightarrow \quad \sqrt{JA^*AJ} = \sum_{\lambda>0} \lambda(\widetilde{P}_\lambda - \widetilde{P}_{-\lambda}), \tag{29}$$

where

$$P_\lambda + P_{-\lambda} = \widetilde{P}_\lambda + \widetilde{P}_{-\lambda} \tag{30}$$

is any decomposition of $P_\lambda + P_{-\lambda}$ into a sum of commuting orthogonal projections $\widetilde{P}_\lambda$, $\widetilde{P}_{-\lambda}$. We have to minimize the distance $d(P,A)$:

$$\langle A-P, A-P\rangle = \langle A,A\rangle - 2\langle A,P\rangle + \langle P,P\rangle \tag{31}$$

Because of Equ. (9) this minimization is equivalent to maximizing $\langle A, P\rangle$:

$$\langle P, A\rangle = \operatorname{tr}(P^{-1}AJ) = \operatorname{tr}(\sqrt{JA^*AJ}) = \sum_{\lambda>0} \lambda(\operatorname{tr}\widetilde{P}_\lambda - \operatorname{tr}\widetilde{P}_{-\lambda}). \tag{32}$$

We see that $\langle P, A\rangle$ is maximal if and only if $\widetilde{P}_{-\lambda} = 0$ for all λ, i.e., if $AJ = P\sqrt[+]{JA^*AJ}$. This (ordinary) polar decomposition is unique precisely for nonsingular A. □

Minimizing the distance to $\mathrm{SO}(V)$

The special orthogonal group in V is denoted by $\mathrm{SO}(V)$. Recall that the polar decomposition $A = PS$ with $P \in \mathrm{O}(V)$ and S symmetric and nonnegative is not unique in the case that A is singular: The set of polar decompositions has the form QPS where $Q \in \mathrm{O}(V)$ leaves the image subspace of A fixed.

Proposition 7 *Assume that $JA^*AJ = \sum_{\lambda>0} \lambda^2 \overline{P}_{\lambda^2}$ is the spectral decomposition, and λ_0^2 is the smallest eigenvalue of JA^*AJ. If $\lambda_0 > 0$, we choose commuting orthogonal projections $\widetilde{P}_{\lambda_0}$, $\widetilde{p}_{-\lambda_0}$ such that*

$$\widetilde{P}_{-\lambda_0} + \widetilde{P}_{\lambda_0} = \overline{P}_{\lambda_0^2}, \quad \operatorname{rk}\widetilde{P}_{-\lambda_0} = 1. \tag{33}$$

The element $P \in \mathrm{SO}(V)$ closest to A is given by

$$\det A > 0: \qquad P = AJ(\sqrt[+]{JA^*AJ})^{-1} \tag{34}$$

$$\det A < 0: \qquad P = AJ\Big(\lambda_0(\widetilde{P}_{\lambda_0} - \widetilde{P}_{-\lambda_0}) + \sum_{\lambda>\lambda_0} \lambda \overline{P}_{\lambda^2}\Big)^{-1} \tag{35}$$

In the case of $\det(A) = 0$, we have to choose P as an orientation-preserving isometry with $AJ = P\sqrt[+]{JA^*AJ}$.

Proof: We use the notation of Prop. 6 and its proof. In the case of nonsingular A Equ. (27) shows that $\det(A) = \det(P)$, so for $\det(A) > 0$ the closest element in $\mathrm{O}(V)$ is already contained in $\mathrm{SO}(V)$. In the case $\det(A) < 0$ we have to choose $\widetilde{P}_{-\lambda}$'s of odd rank such that $\langle A, P\rangle$ is maximal. This is done by selecting the minimal eigenvalue of JA^*AJ and $\operatorname{rk}\widetilde{P}_{-\lambda} = 1$. The case of $\det(A) = 0$ is clear. □

Matrices and the singular value decomposition

In terms of matrices, the polar decomposition is connected to the singular value decomposition: Assume that $V = \mathbb{R}^n$, and that

$$AJ = QDQ' \tag{36}$$

is the singular value decomposition of AJ, with $Q, Q' \in \mathrm{O}_n = \mathrm{O}(\mathbb{R}^n)$, and D a diagonal matrix. Then the polar decomposition is given by

$$AJ = QQ'(Q'^{-1}DQ'). \tag{37}$$

It follows that

$$P = QQ', \quad \sqrt[+]{JA^*AJ} = Q'^{-1}D'. \tag{38}$$

is the element of O_n closest to A.

Medial axis of $\mathrm{O}(V)$ and $\mathrm{SO}(V)$.

The medial axis of a subset X of a metric space Y is defined as the locus of points $y \in Y$ where there are at least two points $x_1, x_2 \in X$ with $d(y, X) = d(y, x_1) = d(y, x_2)$.

Thus the statements of Prop. 6 and Prop. 7 concerning the uniqueness of the closest approximating elements in $\mathrm{O}(V)$ and $\mathrm{SO}(V)$ are reformulated as follows (recall that J was supposed to be nonsingular):

Proposition 8 *The medial axis of* $\mathrm{O}(V)$ *as a subset of* $\mathrm{End}(V)$*, which is endowed with an* L^2 *distance, is precisely the subset of singular linear transformations.*

Proposition 9 *The medial axis of* $\mathrm{SO}(V)$ *as a subset of* $\mathrm{End}(V)$*, which is endowed with an* L^2 *distance, consists of those* A *with* $\det(A) \leq 0$ *with the property that the smallest eigenvalue of* AJ^*AJ *has multiplicity* ≥ 2.

Proof: In the case of $\det(A) < 0$, the choice of $\widetilde{P}_{-\lambda_0}$ in the proof of Prop. 7 is not unique if $\mathrm{rk}\overline{P}_{\lambda_0^2} \geq 2$. As to singular A, if $\mathrm{rk}A = \dim V - 1$, then the multiplicity of 0 as eigenvalue of JA^*AJ equals one, and there is a unique orientation preserving P with $AJ = P\sqrt[+]{JA^*AJ}$. If $\mathrm{rk}A \leq \dim V - 2$, this multiplicity is ≥ 2, and P is no longer unique. $\square$

Analyticity of the approximation

In order to show that the mapping $\mathrm{GL}(V) \to \mathrm{O}(V)$, $A \mapsto AJ(\sqrt[+]{JA^*AJ})^{-1}$ is analytic, we do the following: Choose a positively oriented circle in the right half of the complex plane which encloses the eigenvalues of JA^*AJ. In the right halfplane there are holomorphic functions $\sqrt{\xi}$ and $1/\sqrt{\xi}$ which coincide with the positive square root and its inverse on the positive real axis. Then we compute

$$AJ(\sqrt[+]{JA^*AJ})^{-1} \;=\; \frac{AJ}{2\pi i}\oint_\gamma \frac{1}{\sqrt{\xi}}(\xi I - JA^*AJ)^{-1} d\xi \tag{39}$$

As analyticity is a local property, and γ will enclose the eigenvalues of JA^*AJ for all A in a neighbourhood of A, Equ. (39) describes an analytic function.

The approximation by elements of $\mathrm{SO}(V)$ is analytic away from the set $\det A = 0$ and $\mathrm{SO}(V)$'s medial axis: Consider a negatively oriented circle γ_1 in the right halfplane which encloses the smallest eigenvalue of JA^*AJ, and a positively oriented circle γ_2 which encloses the remaining eigenvalues. Then the approximant in $\mathrm{SO}(V)$ is given by

$$\frac{AJ}{2\pi i}\left(\oint_{\gamma_1} + \oint_{\gamma_2}\right)\frac{1}{\sqrt{\xi}}(\xi I - JA^*AJ)^{-1} d\xi \tag{40}$$

2 The general approximation problem

The problem of approximating an L^2 function from V to V by a Euclidean congruence transformation is broken into several smaller subproblems: The approximation by an affine mapping, and the approximation of this affine mapping by a Euclidean congruence transformation. Moreover, the translational parts of both the affine and the Euclidean transformation can be determined in a way independent of the linear parts.

The linear (or affine) space of affine transformations of the form $x \mapsto a + Ax$ is denoted by

$$\mathit{Aff}(V) = \{\alpha \mid \alpha(x) = a + Ax,\ a \in V,\ A \in \mathrm{End}(V)\} \cong V \times \mathrm{End}(V). \tag{41}$$

The affine transformation $x \mapsto a + Ax$ will be denoted by the pair (a, A). We will assume throughout the rest of this paper that $\mathit{Aff}(V) \subset L^2(V, V)$.

Barycenters

We define the total mass $|\mu|$ and the barycenters b_μ, $b_{F,\mu}$ by

$$|\mu| = \int 1\, d\mu(x), \quad b_\mu = \frac{1}{|\mu|}\int x\, d\mu(x), \quad b_{F,\mu} = \frac{1}{|\mu|}\int F(x)\, d\mu(x). \tag{42}$$

Proposition 10 *The integrals in Equ. (42) are well defined and finite, if we assume that $Aff(V) \subset L^2(V,V)$.*

Proof: We use coordinates with respect to an orthonormal basis $e_1,\dots,e_n$. The j-th coordinate of x is given by $\langle e_j, x\rangle$. We consider the mappings $Cx = e_j$, $P_j x = \langle x, e_j\rangle e_j$, and $F_j x = \langle Fx, e_j\rangle e_j$. By assumption, $C, P_j, F \in L^2(V,V)$. Thus we are able to express $|\mu|$ and the coordinates of the barycenters as scalar L^2 scalar products:

$$\begin{aligned}
|\mu| &= \int 1\, d\mu(x) = \|C\|^2 < \infty, && (43)\\
\langle b_\mu, e_j\rangle &= \int \langle x, e_j\rangle\langle e_j, e_j\rangle\, d\mu(x) = \langle P_j, C\rangle; && (44)\\
& \quad \|Fx\|^2 = \sum \|F_j x\|^2 \Longrightarrow \|F_j\|^2 \le \|F\|^2 < \infty, && (45)\\
\langle b_{F,\mu}, e_j\rangle &= \int \langle Fx, e_j\rangle\langle e_j, e_j\rangle\, d\mu(x) = \langle F_j, C\rangle. && (46)
\end{aligned}$$

□

It is well known and also easily verified that the barycenters b_μ and $b_{F,\mu}$ are affinely invariant in the sense that

$$\begin{aligned}
d\mu'(x) = d\mu(x-a) &\Longrightarrow b_{\mu'} = b_\mu - a, && (47)\\
F'(x) = AF(x) + bx &\Longrightarrow b_{F',\mu} = Ab_{F,\mu} + a. && (48)
\end{aligned}$$

Reduction to the linear case

Proposition 11 *For an L^2 function F we consider $F'(x) = F(x - b_\mu) - b_{F,\mu}$. Then the affine (Euclidean) transformation $\alpha = (a, A)$ is closest to F if and only if A is the element of* End(V) (O(V)*, resp.) closest to F' and $\alpha(b_\mu) = b_{F,\mu}$.*

Proof: Clearly the L^2 distance of F and α with respect to μ is equal to the L^2 distance of $F'' : x \mapsto F(x - b_\mu)$ and $\alpha'' : x \mapsto \alpha(x - b_\mu)$ with respect to μ', where $d\mu'(x) = d\mu(x - b_\mu)$. Then

$$\int F' x\, d\mu'(x) = \int (Fx - b_{F,\mu})\, d\mu(x) = 0, \; \int x\, d\mu'(x) = \int (x - b_\mu)\, d\mu(x) = 0. \tag{49}$$

We note that $\alpha''(x) = Ax - Ab_\mu + a$ and that $F'x = F''x - b_{F,\mu}$, and compute

$$\begin{aligned}
\|F - \alpha\|^2_{L^2_\mu} &= \|F'' - \alpha''\|^2_{L^2_{\mu'}} = \int \|(F'x - Ax) + (\underbrace{b_{F,\mu} - Ab_\mu - a}_{=:\, c})\|^2 \, d\mu'(x) \\
&= \|F' - A\|^2_{L^2_{\mu'}} - 2\langle \textstyle\int (F'x - Ax)\, d\mu'(x), c\rangle + \int \|c\|^2 \, d\mu'(x) \\
&= \|F' - A\|^2_{L^2_{\mu'}} + |\mu| \|\|b_{F,\mu} - Ab_\mu - a\|^2.
\end{aligned}$$

This result shows that in order to minimize $d(\alpha, F)$, we have to minimize $d(F', A)$ and $\|b_{F,\mu} - Ab_\mu - a\|$. □

L^2 approximation by linear mappings

As $\mathrm{End}(V)$ is a linear subspace of $L^2(V,V)$, the problem of approximation by linear mappings is rather trivial: Assume that $F \in L^2(V,V)$, and that $E_{00}, \ldots, E_{nn}$ is a basis of $\mathrm{End}(V)$, such that $A \in \mathrm{End}(V)$ can be written in the form $A = \sum_{i,j} a_{ij} E_{ij}$. Then the approximant $A \in \mathrm{End}(V)$ is found by solving the following system of linear equations for the coefficients a_{ij}:

$$\langle F, E_{kl}\rangle = \sum_{i,j} \langle E_{ij}, E_{kl}\rangle a_{ij}. \tag{50}$$

In order to approximate F by an element P of $\mathrm{O}(V)$, we approximate A, with the methods described in Sec. 1.

It turns out that Equ. (50) reduces to a simple matrix equation.

Proposition 12 *Assume that V is coordinatized with respect to an orthonormal basis $(e_1, \ldots, e_n)$. If $\widetilde{F}$ is defined via its coordinate matrix as below, the linear endomorphism A closest to $F \in L^2(V,V)$ is computed via*

$$A \;=\; \widetilde{F} J^{-1}, \quad \textit{with} \quad (\widetilde{f}_{ij}) \int \langle x, e_j\rangle \langle F(x), e_i\rangle \, d\mu(x). \tag{51}$$

Proof: We choose $E_{ij} = \langle e_j, \cdot\rangle e_i$. We assume that the coordinate matrices of A and J are given by (a_{ij}) and (J_{ij}). Then

$$\langle E_{ij}, E_{kl}\rangle \;=\; \int \langle e_j, x\rangle \langle e_l, x\rangle \langle e_i, e_k\rangle \, d\mu(x) = \langle e_j, J e_l\rangle \delta_{ik} = J_{jl}\delta_{ik}, \tag{52}$$

$$\langle F, E_{ij}\rangle \;=\; \int \langle F(x), \langle e_j, x\rangle e_i\rangle \, d\mu(x). \tag{53}$$

Equ. (50) simplifies to

$$\langle F, E_{kl}\rangle = \sum_{i,j} a_{ij} J_{jl} \delta_{ik} = \sum_j a_{kj} J_{jl} \Longrightarrow (\widetilde{f}_{ij}) = AJ. \tag{54}$$

□

Computations

For applications (see [1]), it is convenient to have formulae which apply to the case that $V = \mathbb{R}^n$ and μ is a finite sum of unit point masses at $x_1, \ldots, x_N$. We identify vectors and linear endomorphisms with their respective coordinate n-tuples and matrices, and we assume that the canonical basis in $\mathbb{R}^n$ is orthonormal. The L^2 function we want to approximate is given by $F(x_i) = y_i$. With the symbol 'svd' for the singular value decomposition we have

$$\begin{array}{rclcrcl} J & = & \sum x_i x_i^T & & |\mu| & = & N \\ b_\mu & = & \frac{1}{N}\sum x_i & & b_{F,\mu} & = & \frac{1}{N}\sum y_i \\ \widetilde{F} & = & \sum (y_i - b_{F,\mu})(x_i - b_\mu)^T & & A & = & \widetilde{F} J^{-1} \end{array}$$

$$Q_1 D Q_2 = \mathrm{svd}(AJ) \quad \Longrightarrow \quad Q_1 Q_2 \text{ is closest to } A\text{, and to } F$$

Acknowledgements

This research was carried out as part of the project P13938-MAT which was supported by the Austrian Science Fund (FWF). The author would like to thank Michael Hofer for his implementing and testing the approximation algorithm.

References

[1] Hofer, M., Pottmann, H., Ravani, B.: Subdivision algorithms for motion design based on homologous points. Technical Report No. 88, Institut für Geometrie, TU Wien, December 2001.

[2] Horn, B.: Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A*, Vol. **4**, No. 4 (1987), 629–642.

[3] Ratiu, T.: The motion of the free n-dimensional rigid body. *Indiana Univ. Math. J.* **29** (1980), 609–629.