The grand tour via geodesic interpolation of 2-frames

Daniel Asimov

Computer Sciences Corporation
NASA Ames Research Center
Mountain View, CA 94305-1000

Andreas Buja

Bell Communications Research
Morristown, NJ 07962-1910

ABSTRACT

Grand tours are a class of methods for visualizing multivariate data, or any finite set of points in n-space. The idea is to create an animation of data projections by moving a 2-dimensional projection plane through n-space. The path of planes used in the animation is chosen so that it becomes dense, that is, it comes arbitrarily close to any plane.

One inspiration for the grand tour was the experience of trying to comprehend an abstract sculpture in a museum. One tends to walk around the sculpture, viewing it from many different angles.

A useful class of grand tours is based on the idea of continuously interpolating an infinite sequence of randomly chosen planes. Visiting randomly (more precisely: uniformly) distributed planes guarantees denseness of the interpolating path.

In computer implementations, 2-dimensional orthogonal projections are specified by two 1-dimensional projections which map to the horizontal and vertical screen dimensions, respectively. Hence, a grand tour is specified by a path of pairs of orthonormal projection vectors.

This paper describes an interpolation scheme for smoothly connecting two pairs of orthonormal vectors, and thus for constructing interpolating grand tours. The scheme is optimal in the sense that connecting paths are geodesics in a natural Riemannian geometry.

1.0 Some terminology

We define and discuss a number of key concepts that will be used in this paper.

Definition: A **2-plane** in $\mathbf{R}^n$ is any 2-dimensional linear subspace of $\mathbf{R}^n$.

Note that in our usage, every 2-plane contains the origin $\mathbf{0} \in \mathbf{R}^n$.

Definition: A **2-frame** in $\mathbf{R}^n$ is any ordered pair of orthonormal vectors in $\mathbf{R}^n$.

Note that any 2-frame uniquely determines a 2-plane.

Notation: Suppose $\mathbf{v}$ and $\mathbf{w}$ are orthonormal vectors in $\mathbf{R}^n$. Then the 2-frame determined by $\mathbf{v}$ and $\mathbf{w}$ (in that order) will be denoted by $(\mathbf{v},\mathbf{w})$.

Notation: The 2-plane determined by the 2-frame $F = (\mathbf{v},\mathbf{w})$ will be denoted by span(F) or span$(\mathbf{v},\mathbf{w})$.

Definition: The **Grassmann manifold**, or **Grassmannian,** $G_{2,n}$ of 2-planes in $\mathbf{R}^n$ is the topological space each point of which represents a distinct 2-plane in R^n.

Definition: The **Stiefel manifold** $V_{2,n}$ of 2-frames in $\mathbf{R}^n$ is the topological space each point of which represents a distinct 2-frame in R^n.

Note that $G_{2,n}$ and $V_{2,n}$ are each locally Euclidean of dimensions 2n-4 and 2n-3 respectively, and each is naturally endowed with an intrinsic metric, or distance function (arising from a natural Riemannian metric structure). Thus it makes sense to discuss the distance between any two points of either of these spaces.

Notation: The distance between points x and y of a metric space M will be denoted by d(x, y).

Definition: A subset X of a metric space M is called **dense** if for any point $m \in M$ and any $\delta > 0$, there exists some $x \in X$ such that $d(x, m) < \delta$.

Intuitively, a dense subset is "all over the place." For every point of M, a point of the subset can be found as near as desired.

Definition: A **grand tour implementation** in $\mathbf{R}^n$ is an algorithm[1] for calculating an arbitrarily long sequence $P_1, P_2, \ldots$ of 2-planes in $\mathbf{R}^n$ such that $X = \{P_1, P_2, \ldots\}$ is a dense subset of $G_{2,n}$.

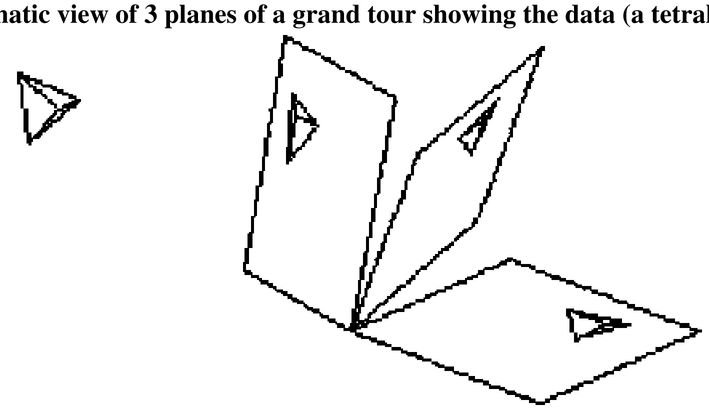
2.0 How grand tours work

The purpose of finding grand tours is to display multivariate data as an animation on a computer screen, allowing the observer to detect patterns in the data. Here is the method that is used. Let $S \subset \mathbf{R}^n$ represent any finite set of multivariate data. Given a grand tour implementation $P_1, P_2, \ldots$, we can now create an animation of S. First, we simply project S orthogonally onto each 2-plane P_i. Now we need some way of identifying the 2-plane P_i with the computer screen. For this purpose, we need to choose, for each i, a 2-frame $F_i = (\mathbf{v}_i, \mathbf{w}_i)$ such that $\mathrm{span}(F_i) = P_i$. Finally, we map the plane P_i to the computer screen via a linear map that takes $\mathbf{v}_i$ and $\mathbf{w}_i$ to the x- and y-directions, respectively, on the computer screen.

The animation now comes from rapidly displaying on the computer screen the result of this procedure for i = 1, 2, ... in turn. At about 10 frames per second, the eye perceives motion; at 24 or more frames per second, the motion is perceived without flicker.

If we have chosen the P_i and the F_i carefully, the result will be a smooth animation of the multivariate data contained in the set S. Each point will appear to be moving on the computer screen along its own smooth trajectory, and points that are actually near each other in S will follow trajectories that remain close to each other throughout the animation. The many different views of S afforded by the dense sequence of 2-planes will offer an observer the opportunity to find patterns in S that may be otherwise difficult to see.

FIGURE 1. Schematic view of 3 planes of a grand tour showing the data (a tetrahedron) projected onto each plane.



3.0 The importance of continuity

Unless we are very careful to make our choices continuous, the animation just described will appear chaotic and will be of no use to a human observer. If either the sequence of 2-planes is not continuous, or the choice of 2-frames is not continuous, then the resulting grand tour implementation will fail to appear continuous. For this reason we refine our original definition of a grand tour implementation as follows:

Definition: A **grand tour** in $\mathbf{R}^n$ is a continuous family of 2-frames { F(t), $0 \le t < \infty$ } such that the corresponding set of 2-planes X = { span(F(t)), $0 \le t < \infty$ } is dense in $G_{2,n}$.

A grand tour will give rise to a grand tour implementation that appears continuous if we choose a sequence of parameter values $0 = t_1 < t_2 < \ldots$ approaching ∞ and sufficiently close to one another. For then the sequence of 2-planes given by P_i = span($F(t_i)$), i = 1, 2, ... will be dense in $G_{2,n}$. In addition, the P_i will appear to vary continuously, and the 2-frames $F(t_i)$ will give us a continuous way to map the image of the data that has been projected onto the planes P_i onto the computer screen.

4.0 Some geometry of the Grassmannian

The aim of this section is to describe geodesic curves in the Grassmann manifold $G_{2,n}$. In the next section these curves will be used to construct a particular grand tour.

In this section P and Q denote any two 2-planes in R^n, or in other words any two points in the Grassmann manifold $G_{2,n}$.

Terminology: A line in a 2-plane is a 1-dimensional linear subspace of the 2-plane. (In other words, it is an ordinary straight line in the 2-plane that passes through the origin of the 2-plane.)

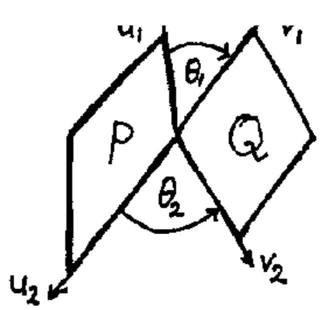
Definition: The first principal angle between two 2-planes P and Q in R^n is the smallest angle between any line in P and any line in Q. The first principal directions between P and Q are the (usually unique) lines in P and Q that realize the first principal angle.

Definition: The second principal directions between two 2-planes P and Q in R^n are the lines in P and Q, respectively, that lie perpendicular to the first principal directions. The second principal angle between P and Q is the smallest angle formed between the second principal directions.

Notation: Thε first and second principal angles will be denoted by θ_1 and θ_2, respectively.

Note that θ_1 and θ_2 will always satisfy the inequalities $0 \le \theta_1 \le \theta_2 \le 90^\circ$.

FIGURE 2. Schematic drawing shows principal angles between two 2-planes in $\mathbf{R}^n$. (In $\mathbf{R}^3$ the first principal angle would necessarily be 0.)



Definition: The distance between P and Q (considered as two points of the Grassmann manifold $G_{2,n}$) will be taken as $d(P, Q) = (\theta_1{}^2 + \theta_2{}^2)^{1/2}$.

Note: Although we are couching this as a definition, in fact this is a consequence of the natural Riemannian metric on $G_{2,n}$.

Definition: A geodesic between any two points p and q of a Riemannian manifold M is a curve in M connecting p to q, which locally minimizes distance.

Suppose that $\alpha : [a,b] \rightarrow M$ is a geodesic with $\alpha(a) = p$ and $\alpha(b) = q$. Suppose we have any sufficiently fine partition of the interval [a, b], say $a = t_0 < t_1 < ... < t_n = b$. Then the definition of geodesic means that for each i = 1,...,n, the curve $\alpha : [t_{i-1}, t_i] \rightarrow M$ will be the shortest of all curves in M connecting $\alpha(t_{i-1})$ to $\alpha(t_i)$.

Notation: Denote by SO_k the group of k x k orthogonal matrices ($A^{-1} = A^t$) with determinant +1.

SO_k is called the special orthogonal group on R^k and is just the group of rotations of R^k.

Definition: Suppose u and v are any two unit vectors in R^n that are an angle θ less than 180° apart. Then slerp(u, v; t) (spherical interpolation) will denote the unique element of SO_n that executes a rotation in span(u, v) by angle tθ, and is the identity on the complementary (n-2)-plane $(\mathrm{span}\,(\mathbf{u}, \mathbf{v}))^{\perp}$.

Fact: Let P and Q be any two 2-planes of R^n. Let $\mathbf{u}_1$ and $\mathbf{v}_1$ denote unit vectors in the first principal directions of P and Q respectively. Let $\mathbf{u}_2$ and $\mathbf{v}_2$ denote unit vectors in the second principal directions of P and Q, respectively. Consider the one-parameter family of SO_n rotations given by the concatenation

$$M(t) = \mathrm{slerp}\,(\mathbf{u}_1, \mathbf{v}_1; t)\,\mathrm{slerp}\,(\mathbf{u}_2, \mathbf{v}_2; t) = \mathrm{slerp}\,(\mathbf{u}_2, \mathbf{v}_2; t)\,\mathrm{slerp}\,(\mathbf{u}_1, \mathbf{v}_1; t)$$

for $0 \le t \le 1$. Then the shortest Grassmannian geodesic[17] between P and Q in $G_{2,n}$ is given by $\alpha : [0, 1] \rightarrow G_{2,n}$ via

$$\alpha(t) = M(t)P$$

Of course, $\alpha(0) = P$ and $\alpha(1) = Q$.

5.0 A Grassmann tour

We review the following method[2,3,9,13] for creating grand tours: Start with an arbitrary 2-plane P in R^n. Pick at random a second arbitrary 2-plane Q in R^n. If $\mathbf{u}_1$, $\mathbf{v}_1$, $\mathbf{u}_2$, $\mathbf{v}_2$ are just as in the previous section, then $M_1(t) = \mathrm{slerp}(\mathbf{u}_1, \mathbf{v}_1; t)\,\mathrm{slerp}(\mathbf{u}_2, \mathbf{v}_2; t)$ defines a continuous 1-parameter family of rotations in SO_4. As above, the 2-planes $M_1(t)P$, for $0 \le t \le 1$, clearly traverse the geodesic from P to Q.

We attempt to iterate this process to obtain a grand tour. Pick a third 2-plane R in $\mathbf{R}^n$ at random. We now do for Q and R the same thing that we just did for P and Q, respectively. Let us denote the resulting one-parameter family of rotations by $M_2(t)$ for $0 \le t \le 1$. Again, we have a geodesic in the Grassmannian between Q and R, given by $M_2(t)Q$ for $0 \le t \le 1$. We then pick a fourth 2-plane, S, at random and continue this procedure indefinitely to obtain a piecewise geodesic curve in $G_{2,n}$ which "connects the dots" (2-planes).

What we really need, though, is a curve of *2-frames* whose underlying 2-planes are precisely the curve we have just described. Fortunately, the 1-parameter families $M_i(t)$ of rotations in $\mathbf{R}^n$ provide the means for this. Since $M_i(0)$ is the identity in SO_n for each i, we may define a grand concatenation M of all the M_i's as follows: For any $s \ge 0$, let $s = n + t$, where n is a positive integer and $0 \le t < 1$. Then we define $M(s) = M_n(t)M_{n-1}(1)M_{n-2}(1)...M_2(1)M_1(1)$. It is easy to verify that this forms a continuous curve of rotations in SO_n, for all real $s \ge 0$. Any 2-frame F fed into this "pipeline" leads to a continuous 1-parameter family $M(t)F$ of 2-frames whose underlying 2-planes, span(M(t)F), form exactly the piecewise geodesic curve of 2-planes described in the previous paragraph. Thus (with probability 1) we will have constructed a grand tour $\{ M(t)F_0 \}$ in this manner.

Note that this method is implemented in XGobi software, available[13] by ftp from Statlib at CMU.

Definition: A grand tour constructed in the fashion described above will be called the **Grassmann tour**.

Notation: Let K denote any subset of M, and let S = $\{s_1, s_2, ...\}$ denote any sequence of points of M. Then #(K,S; n) denotes the number of elements among $\{m_1, m_2, ..., m_n\}$ which lie in the subset K.

Definition: Let M denote a probability space (a measure space whose total measure is 1). We say that a sequence S = $\{s_1, s_2, ...\}$ in M is **well-distributed** in M if for all measurable subsets K of M, the limit as $n \to \infty$ of #(K,S; n)/n is equal to the measure of K.

Intuitively, this just says that the sequence S ultimately visits each part of M in proportion to its measure. A typical application of this concept will be a compact manifold supplied with a measure such that every non-empty open set has positive measure. Note that any well-distributed sequence on such a manifold must automatically be dense.

Note that as a compact Riemannian manifold, the Grassmannian $G_{2,n}$ is naturally a probability space (there is an essentially unique measure, the **uniform distribution**, arising from the intrinsic metric on $G_{2,n}$). One particularly attractive quality of the Grassmann tour described above is that by its construction, it can be shown to be well-distributed in $G_{2,n}$. The advantage of a grand tour's being well-distributed in $G_{2,n}$ is that the views seen on the screen will not prejudice the observer by lingering disproportionately in one or another region of the Grassmannian. Instead they will give a true impression of what different kinds of views of the data are possible.

6.0 From the Grassmann tour to the Stiefel tour

The Grassmann tour satisfies all the abstract requirements for a grand tour: continuity and denseness. In fact, much more holds true: the Grassmann tour is not only continuous but piecewise optimally smooth, a consequence of the piecewise geodesic property. In addition, this tour is not only dense in the Grassmannian, but well-distributed, a consequence of the independent uniform distribution of each of the randomly-selected planes in its construction.

Yet, the Grassmann tour has a shortcoming: it is impossible to prescribe the orientations (i.e., rotational positions) in which the visited planes are seen on the computer screen. This is not a surprise: the Grassmann tour only claims to interpolate 2-planes, not specific 2-frames. This is clear from the construction: if the starting frame in the starting plane P is F, then the ending frame in the target plane Q is M(1)F. This ending frame is uniquely determined by P, Q, and F and is not subject to the viewer's preferences. The construction of M(t) depends, of course, only on P and Q but not on F.

This lack of control over specific 2-frames can be a problem in interactive implementations: grand tours are not very useful unless they are embedded in a set of interactive tools that permit a viewer to manipulate projections generated by a grand tour. Some basic manipulations include storage and retrieval of projections, and revisiting them at any given moment by steering the grand tour back to them (not by backtracking but by direct interpolation). At this point, however, it is essential that each old projection be presented in the screen orientation in which the viewer saw it the first time around; otherwise it may not be recognizable.

At other times, a viewer may wish to visit specific planes, such as the projection onto the first two variables. Since the screen orientation of such a plane cannot be prescribed in the Grassmann tour, a Grassmann geodesic will in all likelihood place the variables in oblique screen orientations on the screen, rather than in the natural horizontal-vertical position.

Thus arises the need for schemes that permit interpolation of a sequence of specific frames rather than just planes. Although the minimum requirements for a grand tour are satisfied by the Grassmann tour, larger implementation and usability issues dictate frame interpolation in some contexts.

Construction of the optimal frame interpolation is the subject of the remainder of this paper. In precise terms, we show how to construct geodesics on the Stiefel manifold $V_{2,n}$ rather than on the Grassmannian $G_{2,n}$.

Stiefel geodesics can be used to construct an interpolating grand tour by choosing a sequence of 2-frames in $\mathbf{R}^n$ independently and at random. Call this sequence $\mathcal{F} = \{F_1, F_2, ...\}$. Now if we could only compute the geodesic of 2-frames in $V_{2,n}$ between each successive pair of 2-frames F_i, F_{i+1} in the sequence, we could "connect the dots" with Stiefel geodesics. Applying this method, then, to each successive pair F_i, F_{i+1} will give us a continuous curve of 2-frames that is piecewise geodesic, threading through each of the F_i's. This grand tour is what we call the Stiefel tour.

7.0 How to connect two 2-frames by Stiefel geodesics

This section will describe how to calculate the geodesic in $V_{2,n}$ between two 2-frames F_1 and F_2 in R^n. (Of necessity, this section will be more technical than the preceding.)

7.1 Reduction to 4 dimensions

First of all, we can simplify matters by restricting attention to the 4-dimensional subspace of R^n that is generated by the two 2-planes $P_1 = \text{span}(F_1)$ and $P_2 = \text{span}(F_2)$. (In the unlikely event that P_1 and P_2 generate only a 3- or 2-dimensional subspace, we may choose any convenient 4-dimensional subspace that contains it.) By an orthogonal change of coordinates, we may assume without loss of generality that this 4-dimensional subspace constitutes the first 4 coordinates of $\mathbf{R}^n$, so we shall call it $\mathbf{R}^4$. By the change of coordinates, our 2-frames F_1 and F_2 must in fact be 2-frames in $\mathbf{R}^4$, i.e., points in $V_{2,4}$. We may assume that this change of coordinates has been chosen so that the 2-frame F_1 consists of the first two standard basis vectors: $F_1 = (\mathbf{e}_1, \mathbf{e}_2)$.

7.2 The relation between geodesics in $V_{2,4}$ and geodesics in SO_4

The mathematical justification for the construction of Stiefel geodesics that we are about to give stems from the theory of Lie groups. Hence we provide some statements that will allow the interested reader to link this material to Lie group theory[11]. There is an important mapping $p : SO_4 \to V_{2,4}$ defined as follows. Let g be any element of SO_4. Then p(g) is the 2-frame $(g(e_1), g(e_2))$. (Intuitively, p maps an orthogonal 4x4 matrix to the 2-frame consisting of its first two columns.) Let $1_2 \times SO_2$ denote the subgroup of SO_4 consisting of those rotations of $\mathbf{R}^4$ which leave fixed the coordinate plane $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$. Then consideration of the mapping p shows that $V_{2,4}$ is in fact a coset manifold, that is, the result of factoring the group SO_4 out by this subgroup $1_2 \times SO_2$. This is expressed by saying that $V_{2,4} = SO_4 / (1_2 \times SO_2)$.

Consequently, SO_4 may be viewed as a principal fibre bundle[12] over the base space $V_{2,4}$ with fibre SO_2.

We can describe the structure of the fibres in greater detail. Let our second 2-frame F_2 be given by the ordered pair of vectors $(\mathbf{u}_1, \mathbf{u}_2)$. Using Gram-Schmidt orthogonalization, we may easily extend this to an ordered 4-tuple of orthonormal vectors, $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4)$. (If we are unlucky, the Gram-Schmidt process gives us a matrix that has determinant -1, but this is easily corrected by replacing the last column with its negative.) We may view these $\mathbf{u}_i$'s as columns of a special orthogonal matrix U such that $p(U) = F_2$. Thus, the matrices of SO_4 which project down to F_2 by applying p are the matrices

$$U_\theta = (\mathbf{u}_1, \mathbf{u}_2, \cos(\theta)\,\mathbf{u}_3 + \sin(\theta)\,\mathbf{u}_4, -\sin(\theta)\,\mathbf{u}_3 + \cos(\theta)\,\mathbf{u}_4)$$

for $0 \le \theta < 2\pi$. The set of these matrices forms one fibre of the fibre bundle, so it is topologically just a circle in SO_4.

The usefulness of SO_4 stems from the fact that any Stiefel geodesic which starts at $F_1 = (\mathbf{e}_1, \mathbf{e}_2)$ and ends at $F_2 = (\mathbf{u}_1, \mathbf{u}_2)$ is the image under p of some SO_4 geodesic that starts at the identity $1_4 = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ and ends at some specific U_θ. The Stiefel geodesic is obtained from that SO_4 that has the shortest length over all values of θ.

We need to find the shortest geodesic α of $V_{2,4}$ which connects the 2-frames F_1 and F_2. According to the previous paragraph, we may determine α by looking at geodesics "upstairs" in SO_4, as follows.

7.3 Geodesics in SO_4 and their normal forms

We need a few facts about SO_4 geodesics[11]. For any square matrix M, denote by exp(M) the matrix exponential, which can be defined in terms of the usual exponential power series.

Fact: Every element of SO_4 is of the form exp(S), where S is a skew-symmetric matrix.

Note that this representation is not unique. Below we will see what the nature of this non-uniqueness is.

Fact: If U = exp(S), S skew-symmetric, then the curve $\alpha(t) = \exp(tS)$ is a geodesic in SO_4 that connects 1_4 and U.

Fact: The length of this geodesic is equal to $\|S\| = \sqrt{\sum_{i=1}^{4}\sum_{j=1}^{4} s_{i,j}^2}$. (Here the $s_{i,j}$ denote the elements of S.)

This is also the so-called Frobenius norm of S. The matrix S is also called the "tangent vector " or "Lie algebra element" for the geodesic $\alpha(t)$ at the starting point $\alpha(0) = 1_4$. In fact, the derivative of $\alpha(t)$ at t = 0 is just S.

Fact: By an orthogonal change of coordinates, any element U of SO_4 may be put in the normal form N, where

$$N = \begin{bmatrix} \cos(K) & -\sin(K) & 0 & 0 \\ \sin(K) & \cos(K) & 0 & 0 \\ 0 & 0 & \cos(L) & -\sin(L) \\ 0 & 0 & \sin(L) & \cos(L) \end{bmatrix}$$

The advantage of this normal form is that it is easy to determine a Lie algebra element S such that exp(S) = N. Namely, S may be chosen as the matrix

$$S = \begin{bmatrix} 0 & -K & 0 & 0 \\ K & 0 & 0 & 0 \\ 0 & 0 & 0 & -L \\ 0 & 0 & L & 0 \end{bmatrix}$$

Obviously, the values of K and L are unique only up to additive multiples of 2π, whence the non-uniqueness of the Lie algebra elements S.

The squared norm of this matrix S is just the sum of the squares of its elements, that is, $2(K^2 + L^2)$. Since $N = A^{-1}UA$ for some orthogonal matrix A, it follows that $U = ANA^{-1} = A\exp(S)A^{-1} = \exp(ASA^{-1})$. Thus $T = ASA^{-1}$ is a Lie algebra element such that exp(T) = U.

Since conjugation by an orthogonal matrix leaves the norm of a matrix unchanged, it follows that the squared norm of T is also $2(K^2 + L^2)$.

7.4 Making sense of the above for the present purpose

The relevance of the above for our problem is that we now know how to cast the problem of finding Stiefel geodesics in terms of the shortest SO_4 geodesic from 1_4 to each of the U_θ's. Our problem at this point is to determine how we can compute the length of these SO_4 geodesics from the original matrix U_θ, and then find the value of θ which minimizes it.

If the above normal form is computed from each θ, one obtains $K = K(\theta)$ and $L = L(\theta)$. Hence also the squared length of a geodesic $2(K^2 + L^2)$ is a function of θ as well. This is the function that we need to minimize.

7.5 The solution

The quantities K and L derived from the normal form are generally not easy to get at, unless one computes these normal forms explicitly. The method for finding a solution to our problem is to circumvent normal forms by getting at easily computable quantities of U_θ that determine the $K(\theta)$ and $L(\theta)$ directly. These quantities are the traces of U_θ and its matrix square. Here is how this works:

Notation: In what follows we shall denote $\cos(K)$ by C_K and $\cos(L)$ by C_L.

Through an easy calculation and application of trigonometric identities, we get that

$$\mathrm{trace}\,(N) = 2(C_K + C_L) \qquad \text{and}$$

$$\mathrm{trace}\,(N^2) = 2(\cos(2K) + \cos(2L)) = 4(C_K^2 + C_L^2 - 1)$$

Since conjugation by an orthogonal matrix leaves the trace unchanged, it follows that $\mathrm{trace}(U_\theta) = 2(C_K + C_L)$ and $\mathrm{trace}((U_\theta)^2) = 4(C_K^{\,2} + C_L^{\,2} - 1)$. Now we set the variable $a = \mathrm{trace}(U_\theta)$ and the variable $b = \mathrm{trace}((U_\theta)^2)$. The resulting two equations and two unknowns boil down to the quadratic equation

$$C^2 - \left(\frac{a}{2}\right)C + (a^2 - b - 4)/8 = 0$$

where C stands for either C_K or C_L. Solving this quadratic equation, we obtain

$$C = \left(a \pm \sqrt{2b - a^2 + 8}\right)/4$$

for C_K and C_L. Hence we may set

$$K = \cos^{-1}\left(\frac{a + \sqrt{2b - a^2 + 8}}{4}\right) \qquad \text{and} \qquad L = \cos^{-1}\left(\frac{a - \sqrt{2b - a^2 + 8}}{4}\right)$$

where $\cos^{-1}$ denotes inverse cosine. Consequently $K^2 + L^2$, which is half of the squared norm $2(K^2 + L^2)$ of U_θ, can be expressed as

$$f(\theta) = \left[\cos^{-1}\left(\frac{a + \sqrt{2b - a^2 + 8}}{4}\right)\right]^2 + \left[\cos^{-1}\left(\frac{a - \sqrt{2b - a^2 + 8}}{4}\right)\right]^2$$

Since $a = \mathrm{trace}(U_\theta)$ and $b = \mathrm{trace}((U_\theta)^2)$ are easily calculated from the matrix U_θ, this expression for $f(\theta)$ may be evaluated numerically for any value of θ. We now numerically minimize $f(\theta)$ over all θ in the range $0 \le \theta < 2\pi$.

Let θ_{min} be the value of θ which minimizes $f(\theta)$. Substituting θ_{min} for θ in the matrix U_θ we get a matrix that we shall call U_{min}. Now we know that the geodesic from 1_4 to U_{min} is the shortest SO_4 geodesic from 1_4 to any element of SO_4 that projects to the second 2-frame, F_2 by the mapping p. As above, we may make use of the normal form N_{min} for U_{min} in order to determine the skew-symmetric matrix T_{min} such that $\exp(T_{min}) = U_{min}$. Finally we can now express the shortest geodesic between the original 2-frames F_1 and F_2 as $\alpha(t) = p(\exp(tT_{min}))$ for $0 \le t \le 1$.

Bibliography

Note: For the reader's convenience, we have chosen to include a number of works relevant to the subject matter, even if they are not directly referred to in the text.

1. Asimov, D., The grand tour: a tool for viewing multidimensional data, SIAM J. Sci. Stat. Computing, **6**, 1, (1985).

2. Buja, A., and Asimov, D. , Grand tour methods: an outline, Proceedings of the 17th Symp. on the Interface Between Comp. Sci. and Stat., Elsevier, 1985.

3. Buja, A., Asimov, D., Hurley, C., and McDonald, J.A., Elements of a viewing pipeline for data analysis, *Dynamic Graphics for Statistics*, W.S. Cleveland and M.E. McGill, eds., Wadsworth, 1988.

4. Buja, A., Hurley, C., and McDonald, J.A., A data viewer for multivariate data, Proceedings of the 18th Symp. on the Interface Between Comp. Sci. and Stat., Elsevier, 1986.

5. Chambers, J.M., Cleveland, W.S., Kleiner, B., and Tukey, P.A., *Graphical methods for data analysis*, Wadsworth, 1983.

6. Cook, D., Buja, A., and Cabrera, J., Grand tour and projection pursuit, Bellcore Technical Memorandum, 1992.

7. Hanson, A.J., and Heng, P.A. , Visualizing a fourth dimension using geometry and light, Proceedings of Visualization '91, IEEE Computer Society Press, 1991.

8. Hurley, C., The data viewer: a program for graphical data analysis, Ph.D. Thesis and Tech. Report, Statistics Dept., University of Washington, Seattle, 1987.

9. Hurley, C., and Buja, A., Analyzing high-dimensional data with motion graphics, SIAM Journal on Scientific and Statistical Computing, **11**, 6 (1990).

10. Littman, M., Swayne, D.F., Dean, N., and Buja, A. , Visualizing the embedding of objects in euclidean space, Proceedings of the 24th Symp. on the Interface Between Comp. Sci. and Stat., 1992.

11. O'Neill, Barrett, *Semi-Riemannian Geometry*, Academic Press, 1983.

12. Steenrod, N., *The Topology of Fibre Bundles*, Princeton University Press, 1951.

13. Swayne, D.F., Cook, D., and Buja, A., XGobi: interactive dynamic graphics in the X window system with a link to S, Proceedings of the Amer. Stat. Assoc. Section on Statistical Graphics, 1991.

14. Tukey, J. and Tukey, P., Graphical display of data sets in 3 or more dimensions, *Interpreting Multivariate Data*, V. Barnett, ed., Wiley, 1981.

15. Wegman, E.J., The grand tour in k dimensions, Tech. report No. 68, Center for Computational Statistics, George Mason University, 1991.

16. Wegman, E.J., and Carr, E.B., Statistical graphics and visualization, Tech. report No. 84, Center for Computational Statistics, George Mason University, 1992.

17. Wong, Y.-C., Differential geometry of Grassmann manifolds, Proc. Nat. Acad. Sci., **57** (1967).

18. Young, F.W., Kent, D.P., and Kuhfeld, W.F., Dynamic graphics for exploring multivariate data, *Dynamic Graphics for Statistics***,** W.S. Cleveland and M.E. McGill, eds., Wadsworth, 1988.

19. Young, F.W., and Rheingans, P., High-dimensional depth-cuing for guided tours of multivariate data, *Computing and Graphics in Statistics*, A. Buja and P.A. Tukey, eds., Springer-Verlag, 1991.