Fitting Flats to Flats
Supplementary Material

Gabriel Dogadov Ugo Finnendahl Marc Alexa
TU Berlin, Computer Graphics Group
cg.tu-berlin.de

1. Riemannian centers in Graff

In the following, we detail the iterative approach (Graff) that we used as a comparison to our method. Recall that the Riemannian center of mass m is defined as a point $\mathcal{P}$ that minimizes the sum of squared geodesic distances to all given points $\{\mathcal{P}_i\}$:

$$m = \underset{\mathcal{P}\in\mathrm{Gr}(k,d)}{\arg\min} \sum_i d^2(\mathcal{P}, \mathcal{P}_i). \tag{1}$$

1.1. Metric

For two k-flats $\mathcal{F}, \mathcal{F}_i$ in $\mathbb{R}^d$, the function

$$d(\mathcal{F}, \mathcal{F}_i) = \left(\sum_{j=1}^{k+1} \phi_j^2\right)^{1/2} \tag{2}$$

with $\{\phi_j\}$ denoting the affine principal angles between $\mathcal{F}$ and $\mathcal{F}_i$, is the standard metric on Graff(k, d) that we implemented. In Sec. 1.3, we explain how this metric can be generalized to flats of different dimensions.

1.2. Gradient computation

Let flats $\mathcal{F}_i$ be given and represented in Stiefel coordinates $\mathbf{Y}_i$. Recall that the gradient of the sum of squared distances in Eq. (1) and Eq. (2) is given by [1]

$$-\sum_{i=1}^{m} \exp_{\mathbf{Y}}^{-1}(\mathbf{Y}_i), \tag{3}$$

for a Stiefel coordinate $\mathbf{Y}$ and $\exp_{\mathbf{Y}}^{-1}(\mathbf{X})$ denoting the derivative of the geodesic that connects $\mathbf{Y}$ and $\mathbf{X}$. This can be computed as [2]

$$\exp_{\mathbf{Y}}^{-1}(\mathbf{Y}_i) = \mathbf{U}_i \tan^{-1}(\mathbf{\Sigma}_i)\mathbf{V}_i^\top, \tag{4}$$

where the matrices on the right-hand side are computed from the SVD

$$(\mathbf{I} - \mathbf{Y}\mathbf{Y}^\top)\mathbf{Y}_i(\mathbf{Y}^\top\mathbf{Y}_i)^{-1} = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^\top. \tag{5}$$

After computing the gradient $\tilde{\mathbf{G}}$ in the ambient space as in Eq. (3) and (4), it has to be projected into the tangent space of $\mathbf{Y}$ by computing $\mathbf{G} = \tilde{\mathbf{G}} - \tilde{\mathbf{G}}\mathbf{Y}\mathbf{Y}^\top$ [2].

1.3. Subspaces of different dimensions

The function in Eq. (2) can be adapted to flats of different dimensions. If $\mathcal{F}$ is a k-flat and $\mathcal{F}_i$ is an l-flat with $k \leq l$, they yield $k+1$ affine principal angles. Note that $\mathbf{Y} \in \mathbb{R}^{(d+1)\times(k+1)}$ and $\mathbf{Y}_i \in \mathbb{R}^{(d+1)\times(l+1)}$. Since the principal angles are by definition as small as possible, the Riemannian distance between $\mathcal{F}$ and $\mathcal{F}_i$ corresponds to the distance between $\mathcal{F}$ and the closest k-flat that is contained in $\mathcal{F}_i$. It also coincides with the distance between $\mathcal{F}_i$ and the closest l-flat that contains $\mathcal{F}$. While this is not a valid distance measure by itself as two distinct flats can have a zero distance (that is, for example, the case when one is a subset of the other), it can be easily re-written to become a valid metric, while yielding the same gradient for their squared function [3]:

$$d_{\mathrm{Graff}}(\mathcal{F}, \mathcal{F}_i) = \left(|k-l|\frac{\pi^2}{4} + \sum_{j=1}^{k+1} \phi_j^2\right)^{1/2}. \tag{6}$$

The function in Eq. (6) corresponds to the distance between $\mathcal{F}$ and the *furthest* k-flat that is a subset of $\mathcal{F}_i$, or equivalently, to the distance between $\mathcal{F}_i$ and the furthest l-flat that contains $\mathcal{F}$. This is achieved by adding additional $|k-l|$ principal angles with the maximum value of $\frac{\pi}{2}$. It is evident that the gradients of the squares of the sums in Eq. (2) and Eq. (6) w.r.t. the Stiefel coordinates of $\mathcal{F}$ coincide, as the term $|k-l|\frac{\pi^2}{4}$ is independent of the orientation and displacement of $\mathcal{F}$.

The derivative of the geodesic in Eq. (4) has the same dimensionality as the Stiefel coordinates of its first argument $\mathbf{Y}$, namely $(d+1)\times(k+1)$. So while the dimension of the flats $\{\mathcal{F}_i\}$ in Eq. (3) may be varying, the gradient has the dimensionality of $\mathbf{Y}$, as desired. The only necessary adjustment is that the matrix $\mathbf{Y}^\top\mathbf{Y}_i$ from Eq. (5) is not square and the inverse has to be replaced by the pseudo-inverse.

1.4. Exponential map

Given a flat $\mathcal{F}$ in Stiefel coordinates $\mathbf{Y}$ and a direction $\mathbf{H}$ that is in the tangent space of $\mathbf{Y}$ (e.g., the gradient of the

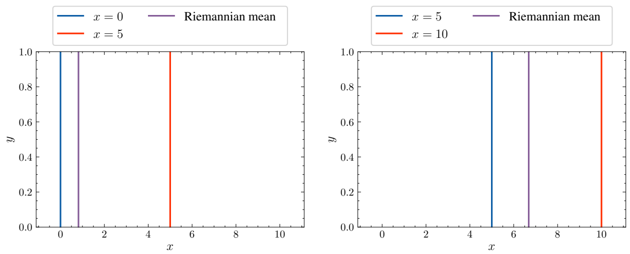

Figure 1. The Riemannian mean (purple) of two parallel lines (blue and red) depends on their distance to the origin.

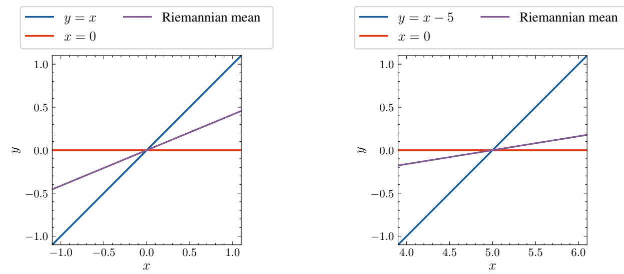

Figure 2. The Riemannian mean (purple) of two intersecting lines (blue and red) depends on the location of their intersection.

geodesic), the image of [2]

$$\mathbf{Y}(t) = (\mathbf{Y}\mathbf{V}\cos(t\boldsymbol{\Sigma}) + \mathbf{U}\sin(t\boldsymbol{\Sigma}))\mathbf{V}^\top$$

yields a geodesic in $\mathrm{Graff}(k,n)$ through $\mathbf{Y}$, with $\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$ being a reduced SVD of $\mathbf{H}$ and $t \in \mathbb{R}$ (e.g., the step size in gradient descent). While the exponential map in Sec. 1.4 yields a point in the Stiefel manifold representing $\mathrm{Gr}(k+1,d+1)$, it may not correspond to a point in $\mathrm{Graff}(k,d)$. Recall that $\mathrm{Graff}(k,d)$ is embedded into $\mathrm{Gr}(k+1,d+1)$ as a submanifold. Thus, we transform the current iterate back into Stiefel coordinates at each iteration.

1.5. Projection into Stiefel coordinates

It is crucial to select a good technique, as it can significantly impact the convergence rate. We tried multiple approaches and found out that one which preserves the column space of $\mathbf{Y}$ works the best, which is detailed in the following.

1. Firstly, it must be ensured that each entry in the last row, except for the last one, is zero. So we compute $m_i = \frac{y_{d+1,k+1}}{y_{d+1,i}}$ for every column $i \in \{1,\dots,k\}$ and then subtract the last column of $\mathbf{Y}$ scaled by m_i from the ith column (this is similar to the Gaussian elimination scheme, but here, we operate on columns instead of rows). After this step, the columns of $\mathbf{Y}$ are no longer orthonormal.

2. Secondly, we orthogonalize the modified $\mathbf{Y}$. In our case, we use the QR decomposition.

1.6. Equivariance properties

Methods that generate Riemann centers on Graff are not equivariant to isometries, i.e., orthogonal transformations and translations – in contrast to the methods we introduce based on squared distance fields. Concretely, Riemannian centers *are* equivariant to orthogonal transformations $\mathbf{R} \in O(d)$, but not to translations.

Orthogonal transformations That the Riemannian center in Graff is equivariant w.r.t. rotations and reflections follows from the fact that affine principal angles, similar to the ordinary principal angles, are invariant. Let $\mathcal{F}_\mathbf{R}$ denote the set of points of the flat $\mathcal{F}$ after $\mathbf{R} \in O(d)$ is applied to them:

$$\mathcal{F}_\mathbf{R} = \{\mathbf{y} \in \mathbb{R}^d \mid \mathbf{y} = \mathbf{R}\mathbf{x},\ \mathbf{x} \in \mathcal{F}\}. \tag{7}$$

Then, $\mathcal{F}_\mathbf{R}$ has the following Stiefel coordinates:

$$\mathbf{Y}_\mathbf{R} = \begin{bmatrix} \mathbf{R}\mathbf{A} & \mathbf{R}\mathbf{b} \\ \mathbf{0}^\top & \sqrt{1+\|\mathbf{b}\|^2} \end{bmatrix}, \tag{8}$$

as $(\mathbf{R}\mathbf{A})^\top\mathbf{R}\mathbf{b} = \mathbf{0}$ if $\mathbf{A}^\top\mathbf{b} = \mathbf{0}$ and $\|\mathbf{R}\mathbf{b}\| = \|\mathbf{b}\|$. It is apparent that the methods on the affine Grassmannian are equivariant, as for any two flats $\mathcal{F},\mathcal{F}'$ it holds that $\mathbf{Y}_\mathbf{R}^\top\mathbf{Y}'_\mathbf{R} = \mathbf{Y}^\top\mathbf{Y}'$, thus the affine principal angles between $\mathcal{F}$ and $\mathcal{F}'$ coincide with the angles between $\mathcal{F}_\mathbf{R}$ and $\mathcal{F}'_\mathbf{R}$. This result holds for any distance measure that depends on the affine principal angles.

Translations Figs. 1 and 2 demonstrate that the Riemannian center in Graff is not equivariant under translations, neither the orientation of the flat nor its shift. Fig. 1 clearly shows that the Riemannian center of two parallel lines is not equidistant to them, and depends on their relative distance to the origin. Fig. 2 illustrates that when two lines intersect in the origin (left), their Riemannian mean's angle is equidistant to the basis lines. If the same lines are shifted away from the origin (right), the mean does not necessarily equal their bisector.

1.7. Alternative method

In the iterative approach on the Grassmannian, since the current iterate is projected into Stiefel coordinates at each iteration anyway, it may be unnecessary to compute the exponential map in some cases. Thus, we have experimented with a variation of the previously described method, which functions as follows: (1) We compute the average of geodesics from $\mathbf{Y}$ to all $\{\mathbf{Y}_i\}$ and add it to $\mathbf{Y}$, and (2) we project the result back onto $\mathrm{Graff}(k,d)$ in Stiefel coordinates. If the input and output dimensions coincide

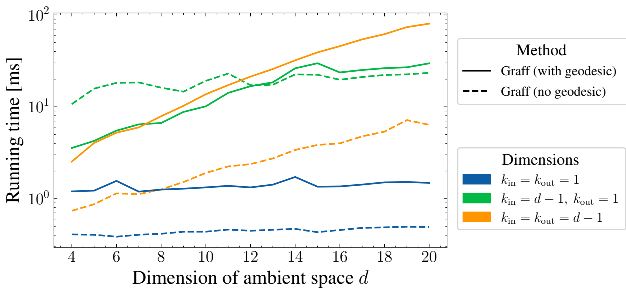

Figure 3. Running time comparison between the iterative approach on the Grassmannian [2] and our modification that omits the computation of the geodesic for different input and output dimension.

($k_{\text{in}} = k_{\text{out}}$), this modification yields similar results as the usual method, while benefitting from a lower running time (c.f. Fig. 3). The lower running time stems from a lower number of iterations and the omission of an SVD calculation per iteration.

References

[1] Hermann Karcher. Riemannian center of mass and so called karcher mean, 2014. 1

[2] Lek-Heng Lim, Ken Sze-Wai Wong, and Ke Ye. Numerical algorithms on the affine grassmannian. *SIAM Journal on Matrix Analysis and Applications*, 40(2):371–393, 2019. 1, 2, 3

[3] Ke Ye and Lek-Heng Lim. Schubert varieties and distances between subspaces of different dimensions. *SIAM Journal on Matrix Analysis and Applications*, 37(3):1176–1197, 2016. 1