Mathematical Formulation of Projection Problem

Jeffrey Liang

February 26, 2025

1 Goal

To project a set of points $\mathcal{X}$ onto the nearest ellipsoid with volume 1, such that the projection of this ellipsoid in two different views produces ellipses which minimise the error to two target ellipses. This error can be measured as the error between the areas of the ellipses or their point clouds.

2 Parametrising Ellipsoids

An ellipsoid $\mathcal{E}$ can be parametrised by a positive definite, symmetric matrix $A \in \mathbb{R}^{3\times 3}$ as follows:

$$\mathcal{E} = \left\{ \mathbf{x} \in \mathbb{R}^3 | \mathbf{x}^\top A \mathbf{x} = 1, A \succ 0 \right\}. \tag{1}$$

By performing an eigen-decomposition on A, we have that A can be represented as $\mathbf{A} = \mathbf{Q}^\top \mathbf{\Lambda} \mathbf{Q}$, where $\mathbf{Q} \in SO(3)$ is an orthonormal rotation matrix and $\mathbf{\Lambda}$ is a diagonal matrix with eigenvalues that are the inverse square of the semi-axes lengths of $\mathcal{E}$:

$$\begin{aligned} A &= \mathbf{Q}^\top \mathbf{\Lambda} \mathbf{Q} && (2) \\ &= \mathbf{Q}^\top \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix} \mathbf{Q} && (3) \\ &= \mathbf{Q}^\top \begin{bmatrix} 1/a^2 & & \\ & 1/b^2 & \\ & & 1/c^2 \end{bmatrix} \mathbf{Q}, && (4) \end{aligned}$$

where a, b, and c are the semi-axes lengths of the ellipsoid $\mathcal{E}$. As the rotation matrix can be parametrised by three Euler angles – roll ψ, pitch θ, and yaw ϕ –

in addition to the three semi-axes, an ellipsoid in 3D can thus be parametrised by six degrees of freedom which we can represent as a vector $\mathbf{u} \in \mathbb{R}^6$.

3 Inner-Level Problem

Let $\mathcal{X} = \{\mathbf{x}_i\}$ be a set of sampled points $\mathbf{x}_i \in \mathbb{R}^3$ from a noisy ellipsoid. Moreover, let $Q : \mathbb{R}^3 \to SO(3)$ map three angles (roll, pitch, yaw) to the orthonormal 3D rotation matrix $\mathbf{Q}$ defined above and $\Lambda : \mathbb{R}^3 \to \mathbb{R}^{3\times 3}$ map three lengths to the diagonal matrix $\mathbf{\Lambda}$ described above.

To project these points to the nearest ellipsoid with a volume of 1 unit cubed, we can minimise the least squares error subject to the volume constraint.

$$\begin{array}{ll} \underset{\text{over } \mathbf{u}\in\mathbb{R}^6}{\text{minimize}} & \sum_{i=1}^{m} \left(\mathbf{x}_i{}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6}) \mathbf{x}_i - 1\right)^2 \\ \text{subject to} & \frac{4}{3}\pi \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 = 1 \end{array} \tag{5}$$

Or, in matrix form, letting

$$\mathbf{X} = \begin{bmatrix} | & \cdots & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_m \\ | & \cdots & | \end{bmatrix},$$

we have

$$\begin{array}{ll} \underset{\text{over } \mathbf{u}\in\mathbb{R}^6}{\text{minimize}} & \left\| \operatorname{diag}(\mathbf{X}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6}) \mathbf{X}) - \mathbf{1}_m \right\|^2 \\ \text{subject to} & \frac{4}{3}\pi \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 = 1 \end{array} \tag{6}$$

4 Outer-Level Problem

There are two approaches to performing the projection of the ellipsoid into two different views: parametrically and via point-sampling.

4.1 Parametric Approach

To perform the projection using parameters obtained from the inner-level problem, we employ orthographic projections. We simulate orthographic projections into two different views by rotating the projected ellipse produced from the inner-level problem differently for each view before projecting the rotated ellipsoids both onto the x-y plane.

Two views were chosen as we believed this was sufficient to constrain the problem such that the ellipsoid matched the desired properties.

The output of the inner-level problem gives us a vector

$$\begin{aligned}\mathbf{u}^* = \quad & \operatorname{argmin}\left\|\operatorname{diag}(\mathbf{X}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6})\mathbf{X}) - \mathbf{1}_m\right\|^2 \\ & \text{subject to} \tfrac{4}{3}\pi \mathbf{u}_1\mathbf{u}_2\mathbf{u}_3 = 1.\end{aligned}$$

Using this $\mathbf{u}^*$, we have the projected ellipsoid with volume 1, and can apply the rotation matrices to the resulting $A(\mathbf{u}^*) = Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6})$. The rotations are performed by two rotation matrices, $\mathbf{R}_1, \mathbf{R}_2 \in SO(3)$ respectively. Each are applied as a right conjugation to the matrix $A(\mathbf{u}^*)$ in a change-of-basis operation.

Thus, we get two rotated ellipsoids, $\mathcal{E}_1$ and $\mathcal{E}_2$, parametrised by the following corresponding matrices:

$$\mathcal{E}_1 : \mathbf{R}_1^\top A(\mathbf{u}^*)\mathbf{R}_1 \qquad \mathcal{E}_2 : \mathbf{R}_2^\top A(\mathbf{u}^*)\mathbf{R}_2. \tag{7}$$

To apply the projection onto the x-y plane, it is *not* as simple as multiplying the projection matrix $\mathbf{P}$,

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{8}$$

as this will give us the cross-section with the x-y plane instead. To obtain the projection, we need to take the Schur complement of the bottom-right entry of the 3D matrix. Let $S : \mathbb{R}^{3\times3} \to \mathbb{R}^{2\times2}$ return the Schur complement described above. Then, the two 2D ellipses ε_1 and ε_2 whose areas we are interested in are parametrised by the matrices $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$:

$$\varepsilon_1 : \hat{\mathbf{A}}_1 = S\big(\mathbf{R}_1^\top A(\mathbf{u}^*)\mathbf{R}_1\big) \qquad \varepsilon_2 : \hat{\mathbf{A}}_2 = S\big(\mathbf{R}_2^\top A(\mathbf{u}^*)\mathbf{R}_2\big)\,. \tag{9}$$

The determinants of these 2×2 matrices $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$ would give the product of the eigenvalues of each matrix. As the eigenvalues would be the squared inverses of the semi-axes of ε_1 and ε_2, the inverse square root of the product of eigenvalues gives the product of the semi-axes lengths. Therefore, the area of an ellipse can be expressed as:

$$\pi ab = \frac{\pi}{\frac{1}{\sqrt{a^2b^2}}} = \frac{\pi}{\sqrt{\lambda_1\lambda_2}} = \frac{\pi}{\sqrt{\det \hat{\mathbf{A}}_i}}. \tag{10}$$

Given two target areas t_1 and t_2 for the projected ellipses ε_1 and ε_2, our final minimisation problem is thus:

$$\begin{array}{ll} \underset{\text{over}\mathcal{X}=\{\mathbf{x}_i\}}{\text{minimize}} & \left\| \frac{\pi}{\sqrt{\det \hat{A}_1(\mathbf{u}^*)}} - t_1 \right\|^2 + \left\| \frac{\pi}{\sqrt{\det \hat{A}_2(\mathbf{u}^*)}} - t_2 \right\|^2 \\ \text{subject to} & \mathbf{u}^* = \text{argmin} \left\| \text{diag}(\mathbf{X}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6})\mathbf{X}) - \mathbf{1}_m \right\|^2 \\ & \quad \text{subject to } \frac{4}{3}\pi \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 = 1 \end{array} .$$

4.2 Sampling Approach

A point-sampling-based approach is more involved, though is more relevant to the water balloon example which this toy example will graduate to. In this approach, we are no longer minimising the error between the ellipses' areas and their targets, but matching the ellipses themselves. As a result, we need to sample the points of the projected *ellipses* and minimise the difference between this set of points and a target set of points (such as a set of points sampled from a circle).

In literature, there have been various proposed metrics to measure the difference between two sets of points. In particular, these metrics are important for point cloud registration, where the goal is to align multiple point clouds. Hausdorff distance has been used previously to evaluate the difference between two point clouds, however, as it is not traditionally differentiable, we will be using Chamfer distance. The Chamfer distance D_C between two sets of points $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$D_C(\mathcal{X}, \mathcal{Y}) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} d(x, \mathcal{Y}) + \frac{1}{|\mathcal{Y}|} \sum_{y \in \mathcal{Y}} d(y, \mathcal{X}), \tag{11}$$

where $d(a, \mathcal{S})$ for a point a and a set $\mathcal{S}$ is defined as $\min d(a, s)$ for all $s \in \mathcal{S}$. In other words, this gives the sum of mean distance between all pairs of points from $\mathcal{X}$ to $\mathcal{Y}$ and the opposite. As Euclidean distance is known to be sensitive to outliers, other variations like density-aware and hyperbolic Chamfer distance have been researched.

Conceptually, after projecting the rotated unit volume ellipsoid onto the x-y plane, we sample around these 2D ellipses. We then calculate the Chamfer distance of these sampled points to our ground truth points at each step and attempt to minimise this distance through backpropagation.

The matter of sampling the points poses issues if not done carefully as stochastic nodes cannot be backpropagated through. We take inspiration from variational autoencoders (VAE) and apply the "reparametrisation trick" such that the random variable can be expressed in terms of deterministic nodes through which the gradient can flow through and a stochastic node to perform the sampling.

Let the projected ellipses ε_1 and ε_2 be represented by the matrices $\hat{\mathbf{A}}_1$ and $\hat{\mathbf{A}}_2$, respectively, as above. For these ellipses to be well-defined, both these matrices must be symmetric positive definite. As proven by Horn & Johnson, there exists a unique positive definite square root $B = A^{1/2}$ for a positive definite and symmetric matrix A such that $A = A^{1/2}A^{1/2} = BB$. Furthermore, as it is positive definite, $A^{1/2}$ is invertible – that is, $A^{-1/2}$ exists.

Thus, the reparametrisation trick applied is to decompose $\hat{\mathbf{A}}_i$ into its square roots like so:

$$\begin{aligned} x^\top \hat{\mathbf{A}}_i x = 1 &\implies x^\top \hat{\mathbf{A}}_i^{1/2} \hat{\mathbf{A}}_i^{1/2} x = 1 && (12) \\ &\implies z^\top z = 1, && (13) \end{aligned}$$

where we let $z = \hat{\mathbf{A}}_i^{1/2} x$. Thus, we simply need to sample unit vectors z (or vectors on the unit circle) and can recover the sampled point x by applying the inverse square root. That is, $x = \hat{\mathbf{A}}^{-1/2} z$.

Let the sampled points of the projected ellipses be denoted by $\mathcal{X}_1(\mathbf{u}^*)$ and $\mathcal{X}_2(\mathbf{u}^*)$ for ε_1 and ε_2, respectively, using the reparametrisation trick. Hence, given two sets of ground truth points $\mathcal{Y}_1, \mathcal{Y}_2$ for the projected ellipses ε_1 and ε_2, our final minimisation problem is

$$\begin{aligned} \underset{\text{over}\mathcal{X}=\{\mathbf{x}_i\}}{\text{minimize}} \quad & d_C(\mathcal{X}_1(\mathbf{u}^*), \mathcal{Y}_1) + d_C(\mathcal{X}_2(\mathbf{u}^*), \mathcal{Y}_2) \\ \text{subject to} \quad & \mathbf{u}^* = \text{argmin} \left\| \text{diag}(\mathbf{X}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6}) \mathbf{X}) - \mathbf{1}_m \right\|^2 \; . \\ & \qquad \text{subject to } \tfrac{4}{3}\pi \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 = 1 \end{aligned}$$

5 Adopting a Mesh-Based Approach

Now, as our water balloon may not remain conform to an ellipsoid, we must represent it more generally. We do this by using a mesh based representation.

Let $\mathcal{M} = (V, E)$ represent the mesh. Let $\text{Vol}(\mathcal{M})$ denote the volume of such a mesh. Then, the first constraint which we must change is $\frac{4}{3}\pi \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 = 1$, which now becomes $\text{Vol}(\mathcal{M}) = 1$.

Next, we replace $\left\| \text{diag}(\mathbf{X}^\top Q(\mathbf{u}_{4:6})^\top \Lambda(\mathbf{u}_{1:3}) Q(\mathbf{u}_{4:6}) \mathbf{X}) - \mathbf{1}_m \right\|^2$ with another least squares function. However, this time, we pair up points V of the mesh with vertices U in the ground-truth mesh we obtain from the physical simulation to form corresponding pairs (v_i, u_i). The replacement least squares function is thus $\sum_{i=1}^{|V|=|U|} \|v_i - u_i\|^2$.

For our upper-level loss function, we can still perform a Chamfer distance to compare with the ground-truth points of the mesh from k different views, $\mathcal{Y}_k$, with these same views of our mesh $\mathcal{X}_k(U)$.

Though this can be confusing, if we take a step back, conceptually, the components of our bi-level optimisation are:

minimize error between projected mesh points V^* and target
subject to projected mesh points V^* = argmin nearest points to existing mesh
subject to volume constraint

This makes it very easy to substitute in our modifications:

$$\begin{array}{ll} \underset{\text{over } \mathcal{X}=\{\mathbf{x}_i\}}{\text{minimize}} & \sum_{k'=1}^{k} d_C(\mathcal{X}_k(V^*), \mathcal{Y}_k) \\ \text{subject to} & V^* = \text{argmin } \sum_{i\in[|\mathcal{X}|]} ||v_i - u_i||^2, \quad v_i \in V, u_i \in U \\ & \qquad \text{subject to Vol}(\mathcal{M}) = 1 \end{array}$$

So the difference between the upper and lower-level losses are that the upper-level function is moving the points $\mathcal{X}$ by minimising differences in 2D projections of the mesh (so only boundary points have gradient flowing through them), whereas the lower-level loss minimises the 3D distance between all points on the mesh with the corresponding pairs.

Additional considerations are now whether the mesh is self-intersecting and ensuring it is watertight; two factors which were implicit in assuming an ellipsoid representation. We need to handle these somehow later on.