Continuous priors on Lie groups

Kristopher Torp Jensen, kris.torp.jensen@gmail.com

Let $g \in \mathcal{G}$ denote an element in group $\mathcal{G}$. Given a trajectory $\{g_1, g_2, ..., g_T\}$ we are interested in its probability under some prior $p(\{g_t\})$ with a notion of smoothness (ideally specified via some learned hyperparameter(s)). In Euclidean space, this is easily achieved by defining a GP prior:

$$p(\{x_t\}) = \mathcal{N}(0, \boldsymbol{K}), \tag{1}$$

where $K_{ij} = k(t_i, t_j)$ for some PSD kernel $k(\cdot,\cdot) : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^n$. Unfortunately this is not easily generalized to non-Euclidean spaces, where we would need $(k(\cdot,\cdot) : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathcal{G})$, since the data on $\mathcal{G}$ cannot be jointly Gaussian (the definition of a GP) as Gaussian distributions are only defined on $\mathbb{R}^n$.

However, we can take inspiration from the fact that the Mátern(p) kernel in Euclidean space is equivalent to a continuous time autoregressive process of order p (AR(p)) for some coefficients $\{a_{1:p}\}$ (Rasmussen & Wiliams appendix B2). Writing such an AR(p) process in discrete time, we get

$$x_t = \epsilon + \sum_{i=1}^{p} a_i x_{t-i}, \tag{2}$$

where $\epsilon \sim \mathcal{N}(0, \Sigma)$. We can re-write this as an AR(p) process over the consecutive *displacements*:

$$(x_t - x_{t-1}) = \epsilon + \sum_{i=1}^{p} a_i (x_{t-i} - x_{t-i-1}) := \delta\tilde{x}_t + \epsilon, \tag{3}$$

which allows us to define a distribution over displacements

$$p(x_t - x_{t-1}) = \mathcal{N}(\delta\tilde{x}_t, \Sigma) \tag{4}$$
$$p((x_t - x_{t-1}) - \delta\tilde{x}_t) = \mathcal{N}(0, \Sigma) \tag{5}$$

We can now write down a prior over trajectories (ignoring boundaries for now), assuming that it factorizes over displacements:

$$p(\{x_t\}) = \prod_t p((x_t - x_{t-1}) - \delta\tilde{x}_t), \tag{6}$$

which provides a basis for a similar definition on $\mathcal{G}$.

On $\mathcal{G}$, we can write the 'displacement' between consecutive states as $(g_{t-1}^{-1} \cdot g_t)$. This gives rise to an 'autoregressive' process

$$g_{t-1}^{-1} \cdot g_t = \epsilon \cdot \prod_{i=1}^{p} a_i (g_{t-i-1}^{-1} \cdot g_{t-i}), \tag{7}$$

where ϵ is a noise process on the manifold (we will define this later). It's not at all clear what we mean by $a_i(g_{t-i-1}^{-1} \cdot g_{t-i})$, which in Euclidean space is a scaled version of the corresponding displacement. Here we will make the following definition, which is equivalent for small displacements:

$$a_i(g_{t-i-1}^{-1} \cdot g_{t-i}) := \text{Exp}\left[a_i \,\text{Log}\left[g_{t-i-1}^{-1} \cdot g_{t-i}\right]\right], \tag{8}$$

where Exp and Log are the exponential and logarithmic maps on $\mathcal{G}$. That is, we take the consecutive displacements, project them into the tangent space, scale them by a learned parameter a_i, and project them back onto the group. Defining

$$\delta\tilde{g}_t := \prod_{i=1}^{p} a_i (g_{t-i-1}^{-1} \cdot g_{t-i}) \tag{9}$$
$$\delta\delta g_t := (\delta\tilde{g}_t)^{-1} \cdot (g_{t-1}^{-1} \cdot g_t), \tag{10}$$

we can now define a distribution over displacements using a Gaussian projected onto the group (as in Falorsi et al. 2019):

$$r(\boldsymbol{x}) = \mathcal{N}(0, \Sigma) \tag{11}$$
$$p(\delta\delta g_t) = \sum_{\boldsymbol{x} \in \mathbb{R}^n :\, \text{Exp}(\boldsymbol{x}) = \delta\delta g_t} r(\boldsymbol{x}) |\boldsymbol{J}(\boldsymbol{x})|^{-1}. \tag{12}$$

We note that for an order 0 process ($p = 0$), $\delta\delta g_t = (g_{t-1}^{-1} \cdot g_t)$ and we recover Brownian motion on the manifold. While this formulation allows us to specify some notion of smoothness/continuity on the manifold, we would like to generalize it to the continuous-time domain as for the Euclidean case, or come up with a different way of specifying a prior over continuous-time smooth processes on such manifolds.