Autonomous Drifting using Simulation-Aided Reinforcement Learning -Additional Material-

Mark Cutler and Jonathan P. How

In this addition to the regular paper, we derive the derivatives needed to use the moment matching method from [1] for doing Gaussian process (GP) predictions when using uncertain inputs. In this case, the GP predictions are combinations of the predictions of two GP's, one from real-world data (rw) and one from simulated data (sim). In the equations below superscript letters encased in parentheses indicate indices of a vector or matrix.

First, since the proportion to which the real data is valued is a function of both the input mean and the covariance (note that $\Sigma_{*_{(rw)}}$ depends on $\boldsymbol{\mu}$ and Σ), the derivatives of $p_{(rw)}$ with respect to $\boldsymbol{\mu}$ and Σ are needed.

The derivative of the generalized logistic function with respect to the input is

$$\frac{\mathrm{d}\,\mathrm{f}(x)}{\mathrm{d}\,x} = -B\,\mathrm{e}^{B(x-x_0)}\left(Q\,\mathrm{e}^{B(x-x_0)}+1\right)^{\left(-\frac{1}{Q}+1\right)}. \tag{1}$$

Defining a helper variable

$$\gamma = \frac{\|\Sigma_{*_{(rw)}}\|_F}{\|\left[\sigma^2_{n_1},\dots,\sigma^2_{n_E}\right]\|},$$

the chain rule of differentiation is applied to get

$$\begin{aligned}\frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}} &= \frac{\mathrm{d}\,\mathrm{f}\,(\gamma)}{\mathrm{d}\,\gamma} \times \frac{\mathrm{d}\,\gamma}{\mathrm{d}\|\Sigma_{*_{(rw)}}\|_F} \times \frac{\mathrm{d}\|\Sigma_{*_{(rw)}}\|_F}{\mathrm{d}\,\Sigma_{*_{(rw)}}} \times \frac{\partial \Sigma_{*_{(rw)}}}{\partial \boldsymbol{\mu}} \\ &= \underbrace{\frac{\mathrm{d}\,\mathrm{f}\,(\gamma)}{\mathrm{d}\,\gamma}}_{1\times1} \times \underbrace{\frac{1}{\|\left[\sigma^2_{n_1},\dots,\sigma^2_{n_E}\right]\|}}_{1\times1} \times \underbrace{\frac{\Sigma_{*_{(rw)}}}{\|\Sigma_{*_{(rw)}}\|_F}}_{E\times E} \times \underbrace{\frac{\partial \Sigma_{*_{(rw)}}}{\partial \boldsymbol{\mu}}}_{E\times E\times D}.\end{aligned}$$

Since $p_{(rw)}$ is a scalar, $\frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}}$ is a $1 \times D$ vector, which is computed element-wise as

$$\frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}^{(i)}} = \frac{\mathrm{d}\,\mathrm{f}\,(\gamma)}{\mathrm{d}\,\gamma} \times \frac{1}{\|\left[\sigma^2_{n_1},\dots,\sigma^2_{n_E}\right]\|} \times \frac{1}{\|\Sigma_{*_{(rw)}}\|_F} \times \sum_{k=1}^{E}\sum_{l=1}^{E} \Sigma^{(k,l)}_{*_{(rw)}} \frac{\partial \Sigma^{(k,l)}_{*_{(rw)}}}{\partial \boldsymbol{\mu}^{(i)}}. \tag{2}$$

Again using the chain rule, the partial derivative of $p_{(rw)}$ with respect to the input covariance is computed as

$$\frac{\partial p_{(rw)}}{\partial \Sigma} = \underbrace{\frac{\mathrm{d}\,\mathrm{f}\,(\gamma)}{\mathrm{d}\,\gamma}}_{1\times1} \times \underbrace{\frac{1}{\|\left[\sigma^2_{n_1},\dots,\sigma^2_{n_E}\right]\|}}_{1\times1} \times \underbrace{\frac{\Sigma_{*_{(rw)}}}{\|\Sigma_{*_{(rw)}}\|_F}}_{E\times E} \times \underbrace{\frac{\partial \Sigma_{*_{(rw)}}}{\partial \Sigma}}_{E\times E\times D\times D},$$

where the $D \times D$ output is calculated element-wise as

$$\frac{\partial p_{(rw)}}{\partial \Sigma^{(i,j)}} = \frac{\mathrm{d}\,\mathrm{f}\,(\gamma)}{\mathrm{d}\,\gamma} \times \frac{1}{\|\left[\sigma^2_{n_1},\dots,\sigma^2_{n_E}\right]\|} \times \frac{1}{\|\Sigma_{*_{(rw)}}\|_F} \times \sum_{k=1}^{E}\sum_{l=1}^{E} \Sigma^{(k,l)}_{*_{(rw)}} \frac{\partial \Sigma^{(k,l)}_{*_{(rw)}}}{\partial \Sigma^{(i,j)}}. \tag{3}$$

Given the derivatives of $p_{(rw)}$, the derivatives of the prediction mean with respect to the input mean can now be calculated as

$$\frac{\partial \mu_*}{\partial \boldsymbol{\mu}} = \frac{\partial \mu_{*_{(sim)}}}{\partial \boldsymbol{\mu}} + p_{(rw)}\left(\frac{\partial \mu_{*_{(rw)}}}{\partial \boldsymbol{\mu}} - \frac{\partial \mu_{*_{(sim)}}}{\partial \boldsymbol{\mu}}\right) + \left(\mu_{*_{(rw)}} - \mu_{*_{(sim)}}\right)\left(\frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}}\right)^T, \tag{4}$$

and the prediction covariance with respect to the input covariance as

$$\frac{\partial \mu^{(i)}_*}{\partial \Sigma} = \frac{\partial \mu^{(i)}_{*_{(sim)}}}{\partial \Sigma} + p_{(rw)}\left(\frac{\partial \mu^{(i)}_{*_{(rw)}}}{\partial \Sigma} - \frac{\partial \mu^{(i)}_{*_{(sim)}}}{\partial \Sigma}\right) + \left(\mu^{(i)}_{*_{(rw)}} - \mu^{(i)}_{*_{(sim)}}\right)\frac{\partial p_{(rw)}}{\partial \Sigma}. \tag{5}$$

Mark Cutler and Jonathan How are with the Laboratory of Information and Decision Systems, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA, USA {cutlerm, jhow}@mit.edu

Similarly, the derivative of the input-output covariance with respect to the input mean is calculated element-wise as

$$\frac{\partial \Sigma_{\boldsymbol{x}_*,f_*}}{\partial \boldsymbol{\mu}^{(i)}} = \frac{\partial \Sigma_{\boldsymbol{x}_*,f_{*(sim)}}}{\partial \boldsymbol{\mu}^{(i)}} + p_{(rw)} \left(\frac{\partial \Sigma_{\boldsymbol{x}_*,f_{*(rw)}}}{\partial \boldsymbol{\mu}^{(i)}} - \frac{\partial \Sigma_{\boldsymbol{x}_*,f_{*(sim)}}}{\partial \boldsymbol{\mu}^{(i)}} \right) + \left(\Sigma_{\boldsymbol{x}_*,f_{*(rw)}} - \Sigma_{\boldsymbol{x}_*,f_{*(sim)}} \right) \frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}^{(i)}}, \tag{6}$$

and with respect to the input covariance as

$$\frac{\partial \Sigma^{(i,j)}_{\boldsymbol{x}_*,f_*}}{\partial \Sigma} = \frac{\partial \Sigma^{(i,j)}_{\boldsymbol{x}_*,f_{*(sim)}}}{\partial \Sigma} + p_{(rw)} \left(\frac{\partial \Sigma^{(i,j)}_{\boldsymbol{x}_*,f_{*(rw)}}}{\partial \Sigma} - \frac{\partial \Sigma^{(i,j)}_{\boldsymbol{x}_*,f_{*(sim)}}}{\partial \Sigma} \right) + \left(\Sigma^{(i,j)}_{\boldsymbol{x}_*,f_{*(rw)}} - \Sigma^{(i,j)}_{\boldsymbol{x}_*,f_{*(sim)}} \right) \frac{\partial p_{(rw)}}{\partial \Sigma}. \tag{7}$$

To calculate the partial derivatives of the predictive covariance, some helper variables are first defined as

$$\Phi^{(i,j)}_{(sim)} = \left(\mu^{(i)}_{*(sim)} - \mu^{(i)}_* \right) \left(\frac{\partial \mu^{(j)}_{*(sim)}}{\partial \boldsymbol{\mu}} - \frac{\partial \mu^{(j)}_*}{\partial \boldsymbol{\mu}} \right) + \left(\mu^{(j)}_{*(sim)} - \mu^{(j)}_* \right) \left(\frac{\partial \mu^{(i)}_{*(sim)}}{\partial \boldsymbol{\mu}} - \frac{\partial \mu^{(i)}_*}{\partial \boldsymbol{\mu}} \right) + \frac{\partial \Sigma^{(i,j)}_{*(sim)}}{\partial \boldsymbol{\mu}}$$

$$\Phi^{(i,j)}_{(rw)} = \left(\mu^{(i)}_{*(rw)} - \mu^{(i)}_* \right) \left(\frac{\partial \mu^{(j)}_{*(rw)}}{\partial \boldsymbol{\mu}} - \frac{\partial \mu^{(j)}_*}{\partial \boldsymbol{\mu}} \right) + \left(\mu^{(j)}_{*(rw)} - \mu^{(j)}_* \right) \left(\frac{\partial \mu^{(i)}_{*(rw)}}{\partial \boldsymbol{\mu}} - \frac{\partial \mu^{(i)}_*}{\partial \boldsymbol{\mu}} \right) + \frac{\partial \Sigma^{(i,j)}_{*(rw)}}{\partial \boldsymbol{\mu}}.$$

Using $\Phi_{(sim)}$ and $\Phi_{(rw)}$, the partial derivative of the predictive covariance with respect to the input mean are calculated as

$$\frac{\partial \Sigma^{(i,j)}_*}{\partial \boldsymbol{\mu}} = \Phi^{(i,j)}_{(sim)} + p_{(rw)} \left(\Phi^{(i,j)}_{(rw)} - \Phi^{(i,j)}_{(sim)} \right) + \left(\beta^{(i,j)}_{(rw)} - \beta^{(i,j)}_{(sim)} \right) \left(\frac{\partial p_{(rw)}}{\partial \boldsymbol{\mu}} \right)^T. \tag{8}$$

Defining similar helper variables for the covariance

$$\Psi^{(i,j)}_{(sim)} = \left(\mu^{(i)}_{*(sim)} - \mu^{(i)}_* \right) \left(\frac{\partial \mu^{(j)}_{*(sim)}}{\partial \Sigma} - \frac{\partial \mu^{(j)}_*}{\partial \Sigma} \right) + \left(\mu^{(j)}_{*(sim)} - \mu^{(j)}_* \right) \left(\frac{\partial \mu^{(i)}_{*(sim)}}{\partial \Sigma} - \frac{\partial \mu^{(i)}_*}{\partial \Sigma} \right) + \frac{\partial \Sigma^{(i,j)}_{*(sim)}}{\partial \Sigma}$$

$$\Psi^{(i,j)}_{(rw)} = \left(\mu^{(i)}_{*(rw)} - \mu^{(i)}_* \right) \left(\frac{\partial \mu^{(j)}_{*(rw)}}{\partial \Sigma} - \frac{\partial \mu^{(j)}_*}{\partial \Sigma} \right) + \left(\mu^{(j)}_{*(rw)} - \mu^{(j)}_* \right) \left(\frac{\partial \mu^{(i)}_{*(rw)}}{\partial \Sigma} - \frac{\partial \mu^{(i)}_*}{\partial \Sigma} \right) + \frac{\partial \Sigma^{(i,j)}_{*(rw)}}{\partial \Sigma},$$

the partial derivative of the predicted covariance with respect to the input covariance are

$$\frac{\partial \Sigma^{(i,j)}_*}{\partial \Sigma} = \Psi^{(i,j)}_{(sim)} + p_{(rw)} \left(\Psi^{(i,j)}_{(rw)} - \Psi^{(i,j)}_{(sim)} \right) + \left(\beta^{(i,j)}_{(rw)} - \beta^{(i,j)}_{(sim)} \right) \frac{\partial p_{(rw)}}{\partial \Sigma}. \tag{9}$$

In summary, Equations 1-9, together with the partial derivatives in [1], [2] define the partial derivatives needed to implement the moment matching algorithm for approximating the output of the combined real-world and simulation GP from the regular paper.

References

[1] M. Deisenroth, D. Fox, and C. Rasmussen, "Gaussian processes for data-efficient learning in robotics and control," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PP, no. 99, 2014.

[2] M. Cutler and J. P. How, "Efficient reinforcement learning for robots using informative simulated priors," in *IEEE International Conference on Robotics and Automation (ICRA)*. Seattle, WA: IEEE, May 2015.