Variational Inference for Inverse Reinforcement Learning with Gaussian Processes: Supplementary Material

Paulius Dilkas (2146879)

6th March 2019

Proof.

1.

$$\begin{aligned}\frac{\partial q(\mathbf{u})}{\partial m} &= q(\mathbf{u})\frac{\partial}{\partial \boldsymbol{\mu}}\left[-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})\right] \\ &= -\frac{1}{2}q(\mathbf{u})(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Sigma}^{-\intercal})(\mathbf{u}-\boldsymbol{\mu})\frac{\partial}{\partial \boldsymbol{\mu}}[\mathbf{u}-\boldsymbol{\mu}] \\ &= \frac{1}{2}q(\mathbf{u})(\boldsymbol{\Sigma}^{-1}+\boldsymbol{\Sigma}^{-\intercal})(\mathbf{u}-\boldsymbol{\mu}).\end{aligned}$$

2. For $\mathbf{S} \in \{\boldsymbol{\Sigma}, \mathbf{B}, \mathbf{D}\}$,

$$\begin{aligned}\frac{\partial q(\mathbf{u})}{\partial \mathbf{S}} &= \frac{\partial}{\partial \mathbf{S}}\left[\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left(-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})\right)\right] \\ &= \frac{\partial}{\partial \mathbf{S}}\left[\frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}}\right]\exp\left(-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})\right) \\ &+ \frac{1}{(2\pi)^{m/2}|\boldsymbol{\Sigma}|^{1/2}}\frac{\partial}{\partial \mathbf{S}}\left[\exp\left(-\frac{1}{2}(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})\right)\right] \\ &= -\frac{1}{2}q(\mathbf{u})\left(|\boldsymbol{\Sigma}|^{-3/2}\frac{\partial|\boldsymbol{\Sigma}|}{\partial \mathbf{S}} + \frac{\partial}{\partial \mathbf{S}}[(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})]\right).\end{aligned} \tag{1}$$

For derivatives with respect to $\boldsymbol{\Sigma}$, we can refer to Petersen and Pedersen [2]:

$$\begin{aligned}\frac{\partial|\boldsymbol{\Sigma}|}{\partial \boldsymbol{\Sigma}} &= |\boldsymbol{\Sigma}|\boldsymbol{\Sigma}^{-\intercal}, \\ \frac{\partial}{\partial \mathbf{S}}[(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})] &= -\boldsymbol{\Sigma}^{-\intercal}\mathbf{U}\boldsymbol{\Sigma}^{-\intercal},\end{aligned} \tag{2}$$

while we can use an online tool by Laue et al.[1] [1] for the remaining ones:

$$\begin{aligned}\frac{\partial|\boldsymbol{\Sigma}|}{\partial \mathbf{B}} &= (\operatorname{adj}(\mathbf{T}) + \operatorname{adj}(\boldsymbol{\Sigma}))\mathbf{B}, \\ \frac{\partial|\boldsymbol{\Sigma}|}{\partial \mathbf{D}} &= \operatorname{adj}(\mathbf{T})\mathbf{D}^\intercal + \mathbf{D}^\intercal \operatorname{adj}(\mathbf{T}), \\ \frac{\partial}{\partial \mathbf{B}}[(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})] &= -(\mathbf{TUT} + \boldsymbol{\Sigma}^{-1}\mathbf{U}\boldsymbol{\Sigma}^{-1})\mathbf{B}, \\ \frac{\partial}{\partial \mathbf{D}}[(\mathbf{u}-\boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1}(\mathbf{u}-\boldsymbol{\mu})] &= -\mathbf{TUTD}^\intercal - \mathbf{D}^\intercal\mathbf{TUT}.\end{aligned} \tag{3}$$

[1] `http://www.matrixcalculus.org/`

Substituting results from (2) and (3) back into (1) gives:

$$\frac{\partial q(\mathbf{u})}{\partial \mathbf{\Sigma}} = \frac{1}{2} q(\mathbf{u})(\mathbf{\Sigma}^{-\intercal}\mathbf{U}\mathbf{\Sigma}^{-\intercal} - \mathbf{\Sigma}^{-\intercal}),$$
$$\frac{\partial q(\mathbf{u})}{\partial \mathbf{B}} = \frac{1}{2} q(\mathbf{u})\{|\mathbf{\Sigma}|^{-3/2}(\mathrm{adj}(\mathbf{T}) + \mathrm{adj}(\mathbf{\Sigma})) + \mathbf{TUT} + \mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{\Sigma}^{-1}\}\mathbf{B},$$
$$\frac{\partial q(\mathbf{u})}{\partial \mathbf{D}} = \frac{1}{2} q(\mathbf{u})\{\mathbf{TUTD}^\intercal + \mathbf{D}^\intercal\mathbf{TUT} - |\mathbf{\Sigma}|^{-3/2}(\mathrm{adj}(\mathbf{T})\mathbf{D}^\intercal + \mathbf{D}^\intercal \,\mathrm{adj}(\mathbf{T}))\}.$$

3. Using a result by Rasmussen and Williams [3],

$$\frac{\partial q(\mathbf{r} \mid \mathbf{u})}{\partial \lambda_i} = q(\mathbf{r} \mid \mathbf{u}) \frac{\partial}{\partial \lambda_i}\left[-\frac{1}{2}\mathbf{u}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u} - \frac{1}{2}\log|\mathbf{K}_{\mathbf{u},\mathbf{u}}|\right] = \frac{1}{2} q(\mathbf{r} \mid \mathbf{u}) \,\mathrm{tr}\left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}\mathbf{u}^\intercal\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-\intercal} - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1})\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right).$$

The remaining derivative is

$$\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i} = \begin{cases} \frac{1}{\lambda_i}\mathbf{K}_{\mathbf{u},\mathbf{u}} & \text{if } i = 0, \\ \mathbf{L}_{\mathbf{u},\mathbf{u}} & \text{otherwise,} \end{cases}$$

where

$$\begin{aligned}
[\mathbf{L}_{\mathbf{u},\mathbf{u}}]_{j,k} &= \frac{\partial}{\partial \lambda_i} k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k}) \\
&= k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k}) \frac{\partial}{\partial \lambda_i}\left[-\frac{1}{2}(\mathbf{x}_{\mathbf{u},j} - \mathbf{x}_{\mathbf{u},k})^\intercal \mathbf{\Lambda}(\mathbf{x}_{\mathbf{u},j} - \mathbf{x}_{\mathbf{u},k}) - \mathbb{1}[j \neq k]\sigma^2 \,\mathrm{tr}(\mathbf{\Lambda})\right] \\
&= k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k}) \frac{\partial}{\partial \lambda_i}\left[-\frac{1}{2}\sum_{l=1}^{d} \lambda_l (x_{\mathbf{u},j,l} - x_{\mathbf{u},k,l})^2 - \mathbb{1}[j \neq k]\sigma^2 \sum_{l=1}^{d} \lambda_l\right] \\
&= k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k}) \left(-\frac{1}{2}(x_{\mathbf{u},j,i} - x_{\mathbf{u},k,i})^2 - \mathbb{1}[j \neq k]\sigma^2\right).
\end{aligned}$$

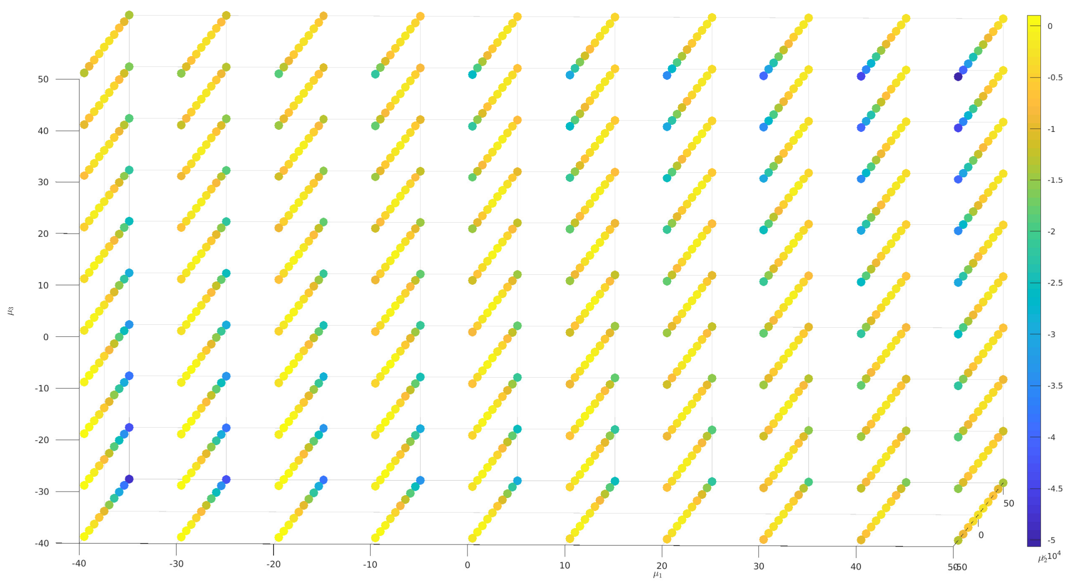
□

References

[1] S. Laue, M. Mitterreiter, and J. Giesen. Computing higher order derivatives of matrix and tensor expressions. In S. Bengio, H. M. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, 3-8 December 2018, Montréal, Canada.*, pages 2755–2764, 2018.

[2] K. B. Petersen, M. S. Pedersen, et al. The matrix cookbook. *Technical University of Denmark*, 7(15):510, 2008.

[3] C. E. Rasmussen and C. K. I. Williams. *Gaussian processes for machine learning.* Adaptive computation and machine learning. MIT Press, 2006.



50
40
30
20
10
0
-10
-20
-30
-40
-40
-30
-20
-10
0
10
20
30
40
50
μ_1
μ_3
μ_2
-50
0
50
0
-0.5
-1
-1.5
-2
-2.5
-3
-3.5
-4
-4.5
-5
10^4