Variational Inference for Inverse Reinforcement Learning with Gaussian Processes

Paulius Dilkas (2146879)

March 4, 2019

ABSTRACT

According to Simon Peyton Jones, an abstract should address four key questions. First, what is the problem that this paper tackles? Second, why is this an interesting problem? Third, what is the solution this paper proposes? Finally, why is the proposed solution a good one?

1. INTRODUCTION

2. BACKGROUND

2.1 Linear Algebra and Numerical Analysis

DEFINITION 2.1 (NORMS). *For any finite-dimensional vector* $\mathbf{x} = (x_1, \dots, x_n)^\intercal$, *its* maximum norm *is*

$$\|\mathbf{x}\|_\infty = \max_i |x_i|$$

whereas its taxicab *(or* Manhattan*)* norm *is*

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|.$$

Let $\mathbf{A}$ *be a matrix. For any vector norm* $\|\cdot\|_p$, *we can also define its* induced norm *for matrices as*

$$\|\mathbf{A}\|_p = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}.$$

In particular, for $p = \infty$, *we have*

$$\|\mathbf{A}\|_\infty = \max_i \sum_j |A_{i,j}|.$$

LEMMA 2.2 (PERTURBATION LEMMA [1]). *Let* $\|\cdot\|$ *be any matrix norm, and let* $\mathbf{A}$ *and* $\mathbf{E}$ *be matrices such that* $\mathbf{A}$ *is invertible and* $\|\mathbf{A}^{-1}\|\|\mathbf{E}\| < 1$, *then* $\mathbf{A}+\mathbf{E}$ *is invertible, and*

$$\|(\mathbf{A}+\mathbf{E})^{-1}\| \le \frac{\|\mathbf{A}^{-1}\|}{1 - \|\mathbf{A}^{-1}\|\|\mathbf{E}\|}.$$

3. THE WIZWOZ SYSTEM

3.1 Notation

For any matrix $\mathbf{A}$, we will use either $A_{i,j}$ or $[\mathbf{A}]_{i,j}$ to denote the element of $\mathbf{A}$ in row i and column j. Moreover, we use $\operatorname{tr}(\mathbf{A})$ to denote its *trace* and $\operatorname{adj}(\mathbf{A})$ for its *adjugate* (or *classical adjoint*).

For any vector $\mathbf{x}$, we write $\mathbb{R}_d[\mathbf{x}]$ to denote a vector space of polynomials with degree at most d, where variables are elements of $\mathbf{x}$, and coefficients are in $\mathbb{R}$.

3.2 Preliminaries

In this paper, all references to measurability are with respect to the Lebesgue measure. Similarly, whenever we consider the existence of an integral, we use the Lebesgue definition of integration.

As recently suggested by Ong et al. [2], we use a decomposition $\mathbf{\Sigma} = \mathbf{B}\mathbf{B}^\intercal + \mathbf{D}^2$, where $\mathbf{B}$ is a lower triangular $m \times p$ matrix with positive diagonal entries, and $\mathbf{D} = \operatorname{diag}(d_1, \dots, d_m)$. Typically, we would set p so that $p \ll m$ to get an efficient approximation, but it is also worth pointing out that we can retain full accuracy by setting $p = m$ and $\mathbf{D} = \mathbf{O}$. Moreover, we define a few matrices that will simplify expressions for the derivatives:

$$\mathbf{U} = (\mathbf{u} - \boldsymbol{\mu})(\mathbf{u} - \boldsymbol{\mu})^\intercal$$
$$\mathbf{S} = \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1},$$
$$\mathbf{\Gamma} = \mathbf{K}_{\mathbf{r},\mathbf{r}} - \mathbf{S}\mathbf{K}_{\mathbf{r},\mathbf{u}},$$
$$\mathbf{R} = \mathbf{S}\frac{\partial \mathbf{K}_{\mathbf{r},\mathbf{u}}}{\partial \lambda_i} - \frac{\partial \mathbf{K}_{\mathbf{r},\mathbf{r}}}{\partial \lambda_i} + \left(\frac{\partial \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal}{\partial \lambda_i} - \mathbf{S}\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i} \right) \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{K}_{\mathbf{r},\mathbf{u}}.$$

Derivatives such as $\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}$ can be expressed as

$$\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i} = \frac{1}{\lambda_i} \mathbf{K}_{\mathbf{u},\mathbf{u}}$$

if $i = 0$, and

$$\left[\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i} \right]_{j,k} = k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k}) \left(-\frac{1}{2}(x_{\mathbf{u},j,i} - x_{\mathbf{u},k,i})^2 - \mathbb{1}[j \neq k]\sigma^2 \right)$$

otherwise.

LEMMA 3.1 (DERIVATIVES OF PROBABILITY DISTRIBUTIONS).

1. $\frac{\partial q(\mathbf{u})}{\partial \boldsymbol{\mu}} = \frac{1}{2} q(\mathbf{u})(\mathbf{\Sigma}^{-1} + \mathbf{\Sigma}^{-\intercal})(\mathbf{u} - \boldsymbol{\mu})$.
2. (a) $\frac{\partial q(\mathbf{u})}{\partial \mathbf{\Sigma}} = \frac{1}{2} q(\mathbf{u})(\mathbf{\Sigma}^{-\intercal}\mathbf{U}\mathbf{\Sigma}^{-\intercal} - \mathbf{\Sigma}^{-\intercal})$.
   (b) $\frac{\partial q(\mathbf{u})}{\partial \mathbf{B}} = q(\mathbf{u})(\mathbf{\Sigma}^{-1}\mathbf{U}\mathbf{\Sigma}^{-1} - |\mathbf{\Sigma}|^{-1}\operatorname{adj}(\mathbf{\Sigma}))\mathbf{B}$.
3. For $i = 0, \dots, d$,
$$\frac{\partial q(\mathbf{r} \mid \mathbf{u})}{\partial \lambda_i} = \frac{1}{2} q(\mathbf{r} \mid \mathbf{u})(|\mathbf{\Gamma}|^{-1}\operatorname{tr}(\mathbf{R}\operatorname{adj}(\mathbf{\Gamma})) - (\mathbf{r} - \mathbf{S}\mathbf{u})^\intercal \mathbf{\Gamma}^{-1}\mathbf{R}\mathbf{\Gamma}^{-1}(\mathbf{r} - \mathbf{S}\mathbf{u})),$$

4. EVALUATION

5. CONCLUSIONS

5.1 Future Work

6. REFERENCES

[1] W. Layton and M. Sussman. *Numerical linear algebra.* Lulu.com, 2014.

[2] V. M.-H. Ong, D. J. Nott, and M. S. Smith. Gaussian variational approximation with a factor covariance structure. *Journal of Computational and Graphical Statistics*, 27(3):465–478, 2018.