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

Paulius Dilkas (2146879)

12th December 2018

1 Preliminaries

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.

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.

Lemma 1.1 (Derivatives of probability distributions)**.**

1. $\frac{\partial q(\mathbf{u})}{\partial \mathbf{m}} = q(\mathbf{u})\frac{1}{2}(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{m})$.
2. $\frac{\partial q(\mathbf{u})}{\partial \mathbf{S}} = -\frac{1}{2}\mathbf{S}^{-\intercal}q(\mathbf{u}) + \frac{1}{2}q(\mathbf{u})\mathbf{S}^{-\intercal}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^\intercal\mathbf{S}^{-\intercal}$.
3. $\frac{\partial q(\mathbf{r})}{\partial \lambda_0} = q(\mathbf{r})\frac{1}{2}\operatorname{tr}\left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}(\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u})^\intercal - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1})\frac{1}{\lambda_0}\mathbf{K}_{\mathbf{u},\mathbf{u}}\right)$.
4. *For* $i = 1, \dots, d$,
$$\frac{\partial q(\mathbf{r})}{\partial \lambda_i} = q(\mathbf{r})\frac{1}{2}\operatorname{tr}\left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}(\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u})^\intercal - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1})\mathbf{L}_{\mathbf{u},\mathbf{u}}\right),$$
where
$$[\mathbf{L}_{\mathbf{u},\mathbf{u}}]_{j,k} = k_{\boldsymbol{\lambda}}(\mathbf{x}_{\mathbf{u},j}, \mathbf{x}_{\mathbf{u},k})\left(-\frac{1}{2}(x_{\mathbf{u},j,i} - x_{\mathbf{u},k,i})^2 - \mathbb{1}[j \neq k]\sigma^2\right).$$

Proof.

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

2.

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

The two remaining derivatives can be taken with the help of *The Matrix Cookbook* [2]:

$$
\begin{aligned}
\frac{\partial}{\partial \mathbf{S}}\left[\frac{1}{|\mathbf{S}|^{1/2}}\right] = -\frac{1}{2}|\mathbf{S}|^{-3/2}\frac{\partial|\mathbf{S}|}{\partial \mathbf{S}} = -\frac{1}{2}|\mathbf{S}|^{-3/2}|\mathbf{S}|\mathbf{S}^{-\intercal} = -\frac{1}{2|\mathbf{S}|^{1/2}}\mathbf{S}^{-\intercal},\\
\frac{\partial}{\partial \mathbf{S}}[(\mathbf{u}-\mathbf{m})^\intercal\mathbf{S}^{-1}(\mathbf{u}-\mathbf{m})] = -\mathbf{S}^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^\intercal\mathbf{S}^{-\intercal}.
\end{aligned}
$$

Plugging them back in gives

$$
\frac{\partial q(\mathbf{u})}{\partial \mathbf{S}} = -\frac{1}{2}\mathbf{S}^{-\intercal}q(\mathbf{u}) + \frac{1}{2}q(\mathbf{u})\mathbf{S}^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^\intercal\mathbf{S}^{-\intercal}.
$$

3.

$$
\frac{\partial q(\mathbf{r})}{\partial \lambda_0} = q(\mathbf{r})\frac{\partial}{\partial \lambda_0}\left[-\frac{1}{2}\mathbf{u}^\intercal\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u} - \frac{1}{2}\log|\mathbf{K}_{\mathbf{u},\mathbf{u}}|\right] = q(\mathbf{r})\frac{1}{2}\operatorname{tr}\left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}(\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u})^\intercal - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1})\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}\right)
$$

by Rasmussen and Williams [3], where

$$
\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0} = \frac{1}{\lambda_0}\mathbf{K}_{\mathbf{u},\mathbf{u}}.
$$

4. The derivation is the same as above, except

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

where

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

□

1.1 Linear Algebra and Numerical Analysis

Definition 1.2 (Norms)**.** For any finite-dimensional vector $\mathbf{x} = (x_1, \ldots, x_n)^\intercal$, 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 an $m \times n$ 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

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

Definition 1.3 (Condition number)**.** For any norm $\|\cdot\|$, the *condition number* of a matrix $\mathbf{A}$ is

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$$

if $\mathbf{A}$ is invertible, and $\kappa(\mathbf{A}) = \infty$ otherwise.

Proposition 1.4.

$$\frac{\|(\mathbf{A} + \mathbf{E})^{-1} - \mathbf{A}^{-1}\|}{\|\mathbf{A}^{-1}\|} \leq \kappa(\mathbf{A}) \frac{\|\mathbf{E}\|}{\|\mathbf{A}\|}$$

2 Proofs

We primarily think of rewards as a vector $\mathbf{r} \in \mathbb{R}^{|\mathcal{S}|}$, but sometimes we use a function notation $r(s)$ to denote the reward of a particular state $s \in \mathcal{S}$. The functional notation is purely a notational convenience.

MDP values are characterised by both a state and a reward function/vector. In order to prove the next theorem, we think of the value function as $V : \mathcal{S} \to \mathbb{R}^{|\mathcal{S}|} \to \mathbb{R}$, i.e., V takes a state $s \in \mathcal{S}$ and returns a function $V(s) : \mathbb{R}^{\mathcal{S}} \to \mathbb{R}$ that takes a reward vector $\mathbf{r} \in \mathbb{R}^{|\mathcal{S}|}$ and returns a value of the state s, $V_{\mathbf{r}}(s) \in \mathbb{R}$. The function $V(s)$ computes the values of all states and returns the value of state s.

Theorem 2.1. *MDP value functions $V(s) : \mathbb{R}^{|\mathcal{S}|} \to \mathbb{R}$ (for $s \in \mathcal{S}$) are Lebesgue measurable.*

Proof sketch. For any reward vector $\mathbf{r} \in \mathbb{R}^{|\mathcal{S}|}$, the collection of converged value functions $\{V_{\mathbf{r}}(s) \mid s \in \mathcal{S}\}$ satisfy

$$\forall s \in \mathcal{S},\ V_{\mathbf{r}}(s) = \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s, a, s') V_{\mathbf{r}}(s')\right).$$

Let $s_i \in \mathcal{S}$ be an arbitrary state. In order to prove that $V(s_i)$ is measurable, it is enough to show that for any $\alpha \in \mathbb{R}$, the set

$$\left\{ \mathbf{r} \in \mathbb{R}^{|\mathcal{S}|} \;\middle|\; \begin{aligned} & V_{\mathbf{r}}(s_i) \in (-\infty, \alpha); \\ & \forall s \in \mathcal{S} \setminus \{s_i\},\ V_{\mathbf{r}}(s) \in \mathbb{R}; \\ & \forall s \in \mathcal{S},\ V_{\mathbf{r}}(s) = \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s, a, s') V_{\mathbf{r}}(s')\right) \end{aligned} \right\}$$

is measurable. Since this set can be constructed in Zermelo-Fraenkel set theory *without* the axiom of choice, it is measurable [1], which proves that $V(s)$ is a measurable function for any $s \in \mathcal{S}$. $\square$

Theorem 2.2. *If the initial values of the MDP value function satisfy the following bound, then the bound remains satisfied throughout value iteration:*

$$|V_{\mathbf{r}}(s)| \leq \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1-\gamma}. \tag{1}$$

Proof. We begin by considering (1) without taking the absolute value of $V_{\mathbf{r}}(s)$, i.e.,

$$V_{\mathbf{r}}(s) \leq \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1-\gamma}, \tag{2}$$

and assuming that the initial values of $\{V_{\mathbf{r}}(s) \mid s \in \mathcal{S}\}$ already satisfy (2). For each $s \in \mathcal{S}$, the value of $V_{\mathbf{r}}(s)$ is updated via this rule:

$$V_{\mathbf{r}}(s) \coloneqq \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s,a,s') V_{\mathbf{r}}(s')\right).$$

Note that both log and exp are increasing functions, $\gamma > 0$, and the $\mathcal{T}$ function gives a probability (a non-negative number). Thus

$$\begin{aligned}
V_{\mathbf{r}}(s) &\leq \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s,a,s') \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1-\gamma}\right) \\
&= \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|)}{1-\gamma} \sum_{s' \in \mathcal{S}} \mathcal{T}(s,a,s')\right) \\
&= \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|)}{1-\gamma}\right)
\end{aligned}$$

by the definition of $\mathcal{T}$. Then

$$\begin{aligned}
V_{\mathbf{r}}(s) &\leq \log\left(|\mathcal{A}| \exp\left(r(s) + \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|)}{1-\gamma}\right)\right) \\
&= \log\left(\exp\left(\log|\mathcal{A}| + r(s) + \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|)}{1-\gamma}\right)\right) \\
&= \log|\mathcal{A}| + r(s) + \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|)}{1-\gamma} \\
&= \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|) + (1-\gamma)(\log|\mathcal{A}| + r(s))}{1-\gamma} \\
&\leq \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|) + (1-\gamma)(\log|\mathcal{A}| + \|\mathbf{r}\|_\infty)}{1-\gamma} \\
&= \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1-\gamma}
\end{aligned}$$

by the definition of $\|\mathbf{r}\|_\infty$.

The proof for

$$V_{\mathbf{r}}(s) \geq \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{\gamma - 1} \tag{3}$$

follows the same argument until we get to

$$\begin{aligned}
V_{\mathbf{r}}(s) &\geq \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|) + (\gamma-1)(\log|\mathcal{A}| + r(s))}{\gamma-1} \\
&\geq \frac{\gamma(\|\mathbf{r}\|_\infty + \log|\mathcal{A}|) + (\gamma-1)(-\log|\mathcal{A}| - \|\mathbf{r}\|_\infty)}{\gamma-1} \\
&= \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{\gamma-1},
\end{aligned}$$

where we use the fact that $r(s) \geq -\|\mathbf{r}\|_\infty - 2\log|\mathcal{A}|$. Combining (2) and (3) gives (1). $\square$

Theorem 2.3 (The Lebesgue Dominated Convergence Theorem [4]). *Let $(X, \mathcal{M}, \mu)$ be a measure space and $\{f_n\}$ a sequence of measurable functions on X for which $\{f_n\} \to f$ pointwise a.e. on X and the function f is measurable. Assume there is a non-negative function g that is integrable over X and dominates the sequence $\{f_n\}$ on X in the sense that*

$$|f_n| \leq g \text{ a.e. on } X \text{ for all } n.$$

Then f is integrable over X and

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proposition 2.4 ([4]). *Let f be a measurable function on E. Suppose there is a non-negative function g that is integrable over E and dominates f in the sense that*

$$|f| \leq g \text{ on } E.$$

Then f is integrable over E.

Theorem 2.5. *Using our usual notation,*

$$\frac{\partial}{\partial t} \iint V_{\mathbf{r}}(s) q(\mathbf{r}) q(\mathbf{u}) \, d\mathbf{r} \, d\mathbf{u} = \iint \frac{\partial}{\partial t} [V_{\mathbf{r}}(s) q(\mathbf{r}) q(\mathbf{u})] \, d\mathbf{r} \, d\mathbf{u},$$

where t is any scalar part of $\mathbf{m}$ or $\mathbf{S}$, or $\boldsymbol{\lambda}$.

Proof. Let

$$f(\mathbf{r}, \mathbf{u}, t) = V_{\mathbf{r}}(s) q(\mathbf{r}) q(\mathbf{u}),$$
$$F(t) = \iint f(\mathbf{r}, \mathbf{u}, t) \, d\mathbf{r} \, d\mathbf{u},$$

and, for any t, let $(t_n)_{n=1}^\infty$ be any sequence such that $\lim_{n\to\infty} t_n = t$, but $t_n \neq t$ for all n. We want to show that

$$F'(t) = \lim_{n\to\infty} \frac{F(t_n) - F(t)}{t_n - t} = \iint \left. \frac{\partial f}{\partial t} \right|_{(\mathbf{r}, \mathbf{u}, t)} d\mathbf{r} \, d\mathbf{u}. \tag{4}$$

We have

$$\frac{F(t_n) - F(t)}{t_n - t} = \iint \frac{f(\mathbf{r}, \mathbf{u}, t_n) - f(\mathbf{r}, \mathbf{u}, t)}{t_n - t} \, d\mathbf{r} \, d\mathbf{u} = \iint f_n(\mathbf{r}, \mathbf{u}) \, d\mathbf{r} \, d\mathbf{u},$$

where

$$f_n(\mathbf{r}, \mathbf{u}) = \frac{f(\mathbf{r}, \mathbf{u}, t_n) - f(\mathbf{r}, \mathbf{u}, t)}{t_n - t}.$$

Since

$$\lim_{n\to\infty} f_n(\mathbf{r}, \mathbf{u}) = \left. \frac{\partial f}{\partial t} \right|_{(\mathbf{r}, \mathbf{u}, t)},$$

(4) follows from Theorem 2.3 as soon as we show that both f and f_n are measurable and find a non-negative integrable function g such that for all n, $\mathbf{r}$, $\mathbf{u}$,

$$|f_n(\mathbf{r}, \mathbf{u})| \leq g(\mathbf{r}, \mathbf{u}).$$

The MDP value function is measurable by Theorem 2.1. The result of multiplying or adding measurable functions (e.g., probability density functions (PDFs)) to a measurable function is still measurable. Thus, both f and f_n are measurable.

It remains to find g. Without loss of generality, assume that t is a parameter of $q(\mathbf{u})$. Then

$$|f_n(\mathbf{r}, \mathbf{u})| = |V_{\mathbf{r}}(s)| q(\mathbf{r}) \left| \frac{q(\mathbf{u})|_{t=t_n} - q(\mathbf{u})}{t_n - t} \right|$$

since PDFs are non-negative. An upper bound for $|V_{\mathbf{r}}(s)|$ is given by Theorem 2.2, while

$$\frac{q(\mathbf{u})|_{t=t_n} - q(\mathbf{u})}{t_n - t} = \left. \frac{\partial q(\mathbf{u})}{\partial t} \right|_{t=c(\mathbf{u})}$$

for some function $c : \mathbb{R}^m \to (\min\{t, t_n\}, \max\{t, t_n\})$ due to the mean value theorem (since q is a continuous and differentiable function of t, regardless of the specific choices of q and t).

Let $\epsilon > 0$ be arbitrary. Then, for sufficiently large n, $|t_n - t| < \epsilon$, and thus

$$|c(\mathbf{u}) - t| < \epsilon. \tag{5}$$

We can rearrange the inequality to produce bounds on $c(\mathbf{u})$ and $|c(\mathbf{u})|$ that will be useful later:

$$\begin{gathered} t - \epsilon < c(\mathbf{u}) < t + \epsilon, \\ |c(\mathbf{u})| < \max\{|t - \epsilon|, |t + \epsilon|\}. \end{gathered} \tag{6}$$

We then have that

$$|f_n(\mathbf{r}, \mathbf{u})| \leq \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r}) \left| \left. \frac{\partial q(\mathbf{u})}{\partial t} \right|_{t=c(\mathbf{u})} \right|.$$

The bound is clearly non-negative and measurable. It remains to show that it is also integrable. Since $\|\mathbf{r}\|_\infty \leq \|\mathbf{r}\|_1$, and

$$\int \frac{\|\mathbf{r}\|_1 + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r})\, d\mathbf{r} = \frac{\log|\mathcal{A}|}{1 - \gamma} + \frac{1}{1 - \gamma} \sum_{i=1}^{k} \mathbb{E}[|r_i|],$$

which clearly exists and is finite, so

$$\int \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r})\, d\mathbf{r} < \infty.$$

Now we just need to show that

$$\iint \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r}) \left| \left. \frac{\partial q(\mathbf{u})}{\partial t} \right|_{t=c(\mathbf{u})} \right| d\mathbf{r}\, d\mathbf{u} \tag{7}$$

exists. Regardless of which of the four derivatives in Lemma 1.1 is taken,

$$\iint \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r}) \left| \left. \frac{\partial q(\mathbf{u})}{\partial t} \right|_{t=c(\mathbf{u})} \right| d\mathbf{r}\, d\mathbf{u} = \iint \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r}) q(\mathbf{u}) |h(\mathbf{u})|\, d\mathbf{r}\, d\mathbf{u},$$

where $h(\mathbf{u})$ is one of:

1. The ith element of the vector
$$\frac{1}{2}(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{c}(\mathbf{u})) \in \mathbb{R}^m,$$
for $i = 1, \ldots, m$, where $\mathbf{c}(\mathbf{u}) = (c_1, \ldots, c_{i-1}, c(\mathbf{u}), c_{i+1} \ldots, c_m)$ (constant everywhere except the ith element).

2. The (i,j)-th element of the matrix

$$-\frac{1}{2}\mathbf{C}(\mathbf{u})^{-\intercal} + \frac{1}{2}\mathbf{C}(\mathbf{u})^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^\intercal\mathbf{C}(\mathbf{u})^{-\intercal} \in \mathbb{R}^{m\times m},$$

for $i,j = 1,\ldots,m$, where $[\mathbf{C}(\mathbf{u})]_{i,j} = c(\mathbf{u})$, and all other elements are constant.

3. $\frac{1}{2}\operatorname{tr}\left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}(\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u})^\intercal - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1})\left.\frac{\partial\mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial\lambda_i}\right|_{\lambda_i=c(\mathbf{u})}\right)$, where $i=0,\ldots,d$.

As

$$\int \frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1-\gamma} q(\mathbf{r})\, d\mathbf{r}$$

gives us a constant, it remains to show that

$$\mathbb{E}_{\mathbf{u}\sim q(\mathbf{u})}[|h(\mathbf{u})|] = \int |h(\mathbf{u})| q(\mathbf{u})\, d\mathbf{u}$$

exists. We will analyse each of the three cases separately.

1. Since

$$\mathbb{E}\left[\frac{1}{2}(\mathbf{S}^{-1}+\mathbf{S}^{-\intercal})(\mathbf{u}-\mathbf{c}(\mathbf{u}))\right] = \frac{1}{2}(\mathbf{S}^{-1}+\mathbf{S}^{-\intercal})\mathbb{E}[\mathbf{u}-\mathbf{c}(\mathbf{u})],$$

we just need to bound $\mathbb{E}[\mathbf{u}-\mathbf{c}(\mathbf{u})]$. For any $j \in \{1,\ldots,m\}\setminus\{i\}$,

$$\mathbb{E}[\mathbf{u}-\mathbf{c}(\mathbf{u})]_j = \mathbb{E}[u_j - c_j] = m_j - c_j,$$

while

$$\mathbb{E}[\mathbf{u}-\mathbf{c}(\mathbf{u})]_i = \mathbb{E}[u_i - c(\mathbf{u})] = m_i - \mathbb{E}[c(\mathbf{u})],$$

where $\mathbb{E}[c(\mathbf{u})]$ exists because $|c(\mathbf{u})|$ has an integrable upper bound in (6).

2. First, we can express $\mathbf{C}(\mathbf{u})$ as $\mathbf{C}(\mathbf{u}) = \mathbf{S}+\mathbf{E}(\mathbf{u})$, where $\mathbf{S}$ is a constant invertible positive semi-definite matrix, and $\mathbf{E} : \mathbb{R}^m \to \mathbb{R}^{m\times m}$ is a matrix-valued function such that $[\mathbf{E}(\mathbf{u})]_{i,j} = c(\mathbf{u}) - t$, while all other elements of $\mathbf{E}(\mathbf{u})$ are zero.

Next, we can divide the problem into two parts; namely, proving the existence of

$$\mathbb{E}[\mathbf{C}(\mathbf{u})^{-\intercal}] \quad \text{and} \quad \mathbb{E}[\mathbf{C}(\mathbf{u})^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^\intercal\mathbf{C}(\mathbf{u})^{-\intercal}].$$

(a) Applying Proposition 1.4 to $\mathbf{S}$ and $\mathbf{E}(\mathbf{u})$ gives

$$\frac{\|\mathbf{C}(\mathbf{u})^{-1} - \mathbf{S}^{-1}\|}{\|\mathbf{S}^{-1}\|} \le \kappa(\mathbf{S})\frac{\|\mathbf{E}(\mathbf{u})\|}{\|\mathbf{S}\|},$$

which can be reformulated to

$$\|\mathbf{C}(\mathbf{u})^{-1} - \mathbf{S}^{-1}\| \le \|\mathbf{S}^{-1}\|^2\|\mathbf{E}(\mathbf{u})\|.$$

Choosing to use the maximum norm we get

$$\max_i \sum_j |[\mathbf{C}(\mathbf{u})^{-1}]_{i,j} - [\mathbf{S}^{-1}]_{i,j}| \le \|\mathbf{S}^{-1}\|^2|c(\mathbf{u}) - t|.$$

Using (5) gives

$$\forall i,\ \sum_j |[\mathbf{C}(\mathbf{u})^{-1}]_{i,j} - [\mathbf{S}^{-1}]_{i,j}| < \|\mathbf{S}^{-1}\|^2\epsilon$$

and

$$\forall i,j,\ |[\mathbf{C}(\mathbf{u})^{-1}]_{i,j} - [\mathbf{S}^{-1}]_{i,j}| < \|\mathbf{S}^{-1}\|^2\epsilon,$$

which bounds all elements of $\mathbf{C}(\mathbf{u})^{-1}$ and proves that $\mathbb{E}[\mathbf{C}(\mathbf{u})^{-\intercal}]$ exists.

(b) Because of the result of 2a, we only need to prove the existence of

$$\mathbb{E}[(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^\intercal].$$

The desired result follows from the existence of $\mathbb{E}[\mathbf{u}]$ and $\mathbb{E}[\mathbf{u}\mathbf{u}^\intercal]$.

3. Again, we can split the proof into two parts: showing the existence of

$$\mathbb{E}\left[\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right|_{\lambda_i=c(\mathbf{u})}\right] \quad \text{and} \quad \mathbb{E}\left[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}\mathbf{u}^\intercal\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-\intercal}\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right|_{\lambda_i=c(\mathbf{u})}\right].$$

(a) In case of $i=0$, each element of $\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}$ is of the form

$$\exp\left(-\frac{1}{2}(\mathbf{x}_j-\mathbf{x}_k)^\intercal\boldsymbol{\Lambda}(\mathbf{x}_j-\mathbf{x}_k)-\mathbb{1}[j\neq k]\sigma^2\operatorname{tr}(\boldsymbol{\Lambda})\right),$$

i.e., without λ_0, so

$$\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}\right|_{\lambda_0=c(\mathbf{u})}=\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}$$

has no $\mathbf{u}$, which means that

$$\mathbb{E}\left[\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}\right|_{\lambda_0=c(\mathbf{u})}\right]=\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}.$$

If $i>0$, then each element of $\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}$ is a constant multiple of $k_{\boldsymbol{\lambda}}(\mathbf{x}_j,\mathbf{x}_k)$ for some $\mathbf{x}_j$ and $\mathbf{x}_k$. Since $k_{\boldsymbol{\lambda}}(\mathbf{x}_j,\mathbf{x}_k)$ is a decreasing function of λ_i, and $c(\mathbf{u})>\lambda_i-\epsilon$,

$$\begin{aligned}
k_{\boldsymbol{\lambda}}(\mathbf{x}_j,\mathbf{x}_k)|_{\lambda_i=c(\mathbf{u})} &= \lambda_0\exp\Bigg(-\frac{1}{2}c(\mathbf{u})(x_{j,i}-x_{k,i})^2-\mathbb{1}[j\neq k]\sigma^2c(\mathbf{u}) \\
&\qquad -\sum_{l\in\{1,\dots,d\}\setminus\{i\}}\frac{1}{2}\lambda_l(x_{j,l}-x_{k,l})^2+\mathbb{1}[j\neq k]\sigma^2\lambda_l\Bigg) \\
&< \lambda_0\exp\Bigg(-\frac{1}{2}(\lambda_i-\epsilon)(x_{j,i}-x_{k,i})^2-\mathbb{1}[j\neq k]\sigma^2(\lambda_i-\epsilon) \\
&\qquad -\sum_{l\in\{1,\dots,d\}\setminus\{i\}}\frac{1}{2}\lambda_l(x_{j,l}-x_{k,l})^2+\mathbb{1}[j\neq k]\sigma^2\lambda_l\Bigg),
\end{aligned}$$

which gives an upper bound on each element of

$$\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right|_{\lambda_i=c(\mathbf{u})}.$$

and shows the existence of

$$\mathbb{E}\left[\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right|_{\lambda_i=c(\mathbf{u})}\right].$$

(b) Since we already found an upper bound for

$$\left.\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i}\right|_{\lambda_i=c(\mathbf{u})}$$

in 3a, $\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] = \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}$, and $\mathbb{E}[\mathbf{u}\mathbf{u}^\intercal]$ clearly exists,

$$\mathbb{E}\left[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{u}\mathbf{u}^\intercal\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-\intercal}\left.\frac{\partial\mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial\lambda_i}\right|_{\lambda_i=c(\mathbf{u})}\right]$$

also exists.

$\square$

Theorem 2.6 (Differentiating under the integral sign)**.** *Assume $f : R \times R \to R$ is such that $x \mapsto f(x,t)$ is measurable for each $t \in R$, that $f(x,t_0)$ is integrable for some $t_0 \in R$ and $\frac{\partial f(x,t)}{\partial t}$ exists for each (x,t). Assume also that there is an integrable $g : R \to R$ with $\left|\frac{\partial f(x,t)}{\partial t}\right| \le g(x)$ for each $x, t \in R$. Then the function $x \mapsto f(x,t)$ is integrable for each t and the function $F : R \to R$ defined by*

$$F(t) = \int_R f_t\, d\mu = \int_R f(x,t)\, d\mu(x)$$

is differentiable with derivative

$$F'(t) = \frac{d}{dt}\int_R f(x,t)\, d\mu(x) = \int_R \frac{\partial}{\partial t} f(x,t)\, d\mu(x).$$

Proof. To prove the formula for $F(t)$ consider any sequence $(t_n)_{n=1}^\infty$ so that $\lim_{n\to\infty} t_n = t$ but $t_n \neq t$ for each t. We claim that

$$\lim_{n\to\infty}\frac{F(t_n) - F(t)}{t_n - t} = \int_R \frac{\partial f(x,t)}{\partial t}\, d\mu(x). \tag{8}$$

We have

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_R \frac{f(x,t_n) - f(x,t)}{t_n - t}\, d\mu(x) = \int_R f_n(x)\, d\mu(x)$$

where

$$f_n(x) = \frac{f(x,t_n) - f(x,t)}{t_n - t}.$$

Notice that, for each x we know

$$\lim_{n\to\infty} f_n(x) = \left.\frac{\partial f}{\partial t}\right|_{(x,t)}$$

and so (8) will follow from the dominated convergence theorem once we show that $|f_n(x)| \le g(x)$ for each x.

That follows from the mean value theorem again because there is c between t and t_0 (with c depending on x) so that

$$f_n(x) = \frac{f(x,t_n) - f(x,t)}{t_n - t} = \left.\frac{\partial f}{\partial t}\right|_{(x,c)}.$$

So $|f_n(x)| \le g(x)$ for each x. $\square$

3 Derivatives of the Evidence Lower Bound

3.1 $\partial/\partial\mathbf{m}$

We begin by removing terms independent of $\mathbf{m}$:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\mathbf{m}} = &-\frac{1}{2}\frac{\partial}{\partial\mathbf{m}}[\mathbf{m}^\intercal\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{m}] + \frac{\partial}{\partial\mathbf{m}}[\mathbf{t}^\intercal\mathbb{E}[\mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{m}] \\ &- \sum_{i=1}^N\sum_{t=1}^T\frac{\partial}{\partial\mathbf{m}}\mathbb{E}[V_\mathbf{r}(s_{i,t})] - \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t},a_{i,t},s')\frac{\partial}{\partial\mathbf{m}}\mathbb{E}[V_\mathbf{r}(s')].\end{aligned}$$

Here

$$\frac{\partial}{\partial \mathbf{m}}[\mathbf{m}^\intercal \mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{m}] = (\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] + \mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^\intercal)\mathbf{m},$$

$$\frac{\partial}{\partial \mathbf{m}}[\mathbf{t}^\intercal \mathbb{E}[\mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{m}] = \mathbf{t}^\intercal \mathbb{E}[\mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}],$$

and

$$\begin{aligned}\frac{\partial}{\partial \mathbf{m}}\mathbb{E}[V_{\mathbf{r}}(s)] &= \frac{\partial}{\partial \mathbf{m}}\iiint V_{\mathbf{r}}(s)p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X}_{\mathbf{u}}, \mathbf{u})\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda})\,d\mathbf{r}\,d\mathbf{u}\,d\boldsymbol{\lambda} \\ &= \iiint V_{\mathbf{r}}(s)p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X}_{\mathbf{u}}, \mathbf{u})\frac{\partial}{\partial \mathbf{m}}\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda})\,d\mathbf{r}\,d\mathbf{u}\,d\boldsymbol{\lambda},\end{aligned} \tag{9}$$

where Substituting it back into (9) gives

$$\begin{aligned}\frac{\partial}{\partial \mathbf{m}}\mathbb{E}[V_{\mathbf{r}}(s)] &= \frac{1}{2}\iiint V_{\mathbf{r}}(s)(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{m})p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X}_{\mathbf{u}}, \mathbf{u})\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda})\,d\mathbf{r}\,d\mathbf{u}\,d\boldsymbol{\lambda} \\ &= \frac{1}{2}\mathbb{E}[V_{\mathbf{r}}(s)(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{m})].\end{aligned}$$

Hence

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{m}} = &- \frac{1}{2}(\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] + \mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^\intercal)\mathbf{m} + \mathbf{t}^\intercal \mathbb{E}[\mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] \\ &- \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}[V_{\mathbf{r}}(s_{i,t})(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{m})] \\ &- \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t}, a_{i,t}, s')\mathbb{E}[V_{\mathbf{r}}(s')(\mathbf{S}^{-1} + \mathbf{S}^{-\intercal})(\mathbf{u} - \mathbf{m})].\end{aligned}$$

3.2 $\partial/\partial\mathbf{S}$

Similarly to the previous section,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \mathbf{S}} = &\frac{1}{2}\frac{\partial}{\partial \mathbf{S}}\log|\mathbf{S}| - \frac{1}{2}\frac{\partial}{\partial \mathbf{S}}\operatorname{tr}[\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{S}] \\ &- \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\partial}{\partial \mathbf{S}}\mathbb{E}[V_{\mathbf{r}}(s_{i,t})] - \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t}, a_{i,t}, s')\frac{\partial}{\partial \mathbf{S}}\mathbb{E}[V_{\mathbf{r}}(s')],\end{aligned}$$

where

$$\frac{\partial}{\partial \mathbf{S}}\log|\mathbf{S}| = \mathbf{S}^{-\intercal},$$

and

$$\frac{\partial}{\partial \mathbf{S}}\operatorname{tr}[\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]\mathbf{S}] = \mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^\intercal$$

by *The Matrix Cookbook* [2]. Then

$$\frac{\partial}{\partial \mathbf{S}}\mathbb{E}[V_{\mathbf{r}}(s)] = \iiint V_{\mathbf{r}}(s)q(\mathbf{r})\frac{\partial}{\partial \mathbf{S}}\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda})\,d\mathbf{r}\,d\mathbf{u}\,d\boldsymbol{\lambda},$$

where and

$$\begin{aligned}\frac{\partial}{\partial \mathbf{S}}\mathbb{E}[V_{\mathbf{r}}(s)] &= \frac{1}{2}\iiint V_{\mathbf{r}}(s)(\mathbf{S}^{-\intercal}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^\intercal\mathbf{S}^{-\intercal} - \mathbf{S}^{-\intercal})q(\mathbf{r})\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda})\,d\mathbf{r}\,d\mathbf{u}\,d\boldsymbol{\lambda} \\ &= \frac{1}{2}\mathbb{E}[V_{\mathbf{r}}(s)(\mathbf{S}^{-\intercal}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^\intercal\mathbf{S}^{-\intercal} - \mathbf{S}^{-\intercal})].\end{aligned}$$

Therefore

$$\frac{\partial \mathcal{L}}{\partial \mathbf{S}} = \frac{1}{2}\mathbf{S}^{-\intercal} - \frac{1}{2}\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^{\intercal} - \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}[V_{\mathbf{r}}(s_{i,t})(\mathbf{S}^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^{\intercal}\mathbf{S}^{-\intercal} - \mathbf{S}^{-\intercal})]$$
$$- \gamma \sum_{s'\in\mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s')\mathbb{E}[V_{\mathbf{r}}(s')(\mathbf{S}^{-\intercal}(\mathbf{u}-\mathbf{m})(\mathbf{u}-\mathbf{m})^{\intercal}\mathbf{S}^{-\intercal} - \mathbf{S}^{-\intercal})].$$

3.3 $\partial/\partial\alpha_j$

We begin in the usual way:

$$\frac{\partial \mathcal{L}}{\partial \alpha_j} = -\frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathbf{m}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\log|\mathbf{K}_{\mathbf{u},\mathbf{u}}|]$$
$$+ \frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathbf{t}^{\intercal}\mathbf{K}_{\mathbf{r},\mathbf{u}}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}] + \frac{\partial}{\partial \alpha_j}[\alpha_j + \log\Gamma(\alpha_j) + (1-\alpha_j)\psi(\alpha_j)]$$
$$- \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\partial}{\partial \alpha_j}\mathbb{E}[V_{\mathbf{r}}(s_{i,t})] - \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t}, a_{i,t}, s')\frac{\partial}{\partial \alpha_j}\mathbb{E}[V_{\mathbf{r}}(s')].$$

First,

$$\frac{\partial}{\partial \alpha_j}[\alpha_j + \log\Gamma(\alpha_j) + (1-\alpha_j)\psi(\alpha_j)] = 1 + \psi(\alpha_j) - \psi(\alpha_j) + (1-\alpha_j)\psi'(\alpha_j) = 1 + (1-\alpha_j)\psi'(\alpha_j)$$

by the definition of ψ. The remaining terms can all be treated in the same way, as they all contain expectations of scalar functions that are independent of α_j, and α_j only occurs in $\Gamma(\lambda_j; \alpha_j, \beta_j)$. Thus we can work with an abstract function as follows:

$$\frac{\partial}{\partial \alpha_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] = \frac{\partial}{\partial \alpha_j}\iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r})q(\boldsymbol{\lambda})q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}$$
$$= \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r})q(\lambda_0)\cdots q(\lambda_{j-1})\frac{\partial}{\partial \alpha_j}\left[\frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)}\lambda_j^{\alpha_j-1}\right]e^{-\beta_j\lambda_j}$$
$$q(\lambda_{j+1})\cdots q(\lambda_d)q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}.$$

Then

$$\frac{\partial}{\partial \alpha_j}\left[\frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)}\lambda_j^{\alpha_j-1}\right] = \frac{\frac{\partial}{\partial \alpha_j}[\beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}]\Gamma(\alpha_j) - \beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}\Gamma'(\alpha_j)}{(\Gamma(\alpha_j))^2}$$
$$= \frac{\beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}\frac{\partial}{\partial \alpha_j}[\alpha_j\log\beta_j + (\alpha_j-1)\log\lambda_j]\Gamma(\alpha_j) - \beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}\Gamma'(\alpha_j)}{(\Gamma(\alpha_j))^2}$$
$$= \frac{\beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}(\log\beta_j + \log\lambda_i)\Gamma(\alpha_j) - \beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}\Gamma'(\alpha_j)}{(\Gamma(\alpha_j))^2}$$
$$= \frac{\beta_j^{\alpha_j}\lambda_j^{\alpha_j-1}}{\Gamma(\alpha_j)}\left(\log\beta_j + \log\lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)}\right),$$

which means that

$$\begin{aligned}
\frac{\partial}{\partial \alpha_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] = & \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) q(\lambda_0) \cdots q(\lambda_{j-1}) \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \lambda_j^{\alpha_j - 1} e^{-\beta_j \lambda_j} \left(\log \beta_j + \log \lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \\
& \qquad q(\lambda_{j+1}) \cdots q(\lambda_d) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u} \\
&= \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\log \beta_j + \log \lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) q(\boldsymbol{\lambda}) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u} \\
&= \mathbb{E}\left[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\log \beta_j + \log \lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \right] \\
&= \left(\log \beta_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] + \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \log \lambda_j].
\end{aligned}$$

With these results in mind, we can simplify the initial expression to

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \alpha_j} &= 1 + (1 - \alpha_j)\psi'(\alpha_j) + \left(\log \beta_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \Bigg(-\frac{1}{2}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]] \\
&- \frac{1}{2}\mathbb{E}[\mathbf{m}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}] - \frac{1}{2}\mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}|] + \mathbb{E}[\mathbf{t}^\intercal \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}] \\
&- \sum_{i=1}^N \sum_{t=1}^T \mathbb{E}[V_{\mathbf{r}}(s_{i,t})] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \mathbb{E}[V_{\mathbf{r}}(s')] \Bigg) \\
&- \frac{1}{2}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}] \log \lambda_j] - \frac{1}{2}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}] \log \lambda_j] - \frac{1}{2}\mathbb{E}[\mathbf{m}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \log \lambda_j] \\
&- \frac{1}{2}\mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}| \log \lambda_j] + \mathbb{E}[\mathbf{t}^\intercal \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \log \lambda_j] \\
&- \sum_{i=1}^N \sum_{t=1}^T \mathbb{E}[V_{\mathbf{r}}(s_{i,t}) \log \lambda_j] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \mathbb{E}[V_{\mathbf{r}}(s') \log \lambda_j].
\end{aligned}$$

3.4 $\partial / \partial \beta_j$

Finally,

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \beta_j} &= -\frac{1}{2}\frac{\partial}{\partial \beta_j}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2}\frac{\partial}{\partial \beta_j}\mathbb{E}[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]] - \frac{1}{2}\frac{\partial}{\partial \beta_j}\mathbb{E}[\mathbf{m}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}] \\
&- \frac{1}{2}\frac{\partial}{\partial \beta_j}\mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}|] + \frac{\partial}{\partial \beta_j}\mathbb{E}[\mathbf{t}^\intercal \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}] - \frac{\partial}{\partial \beta_j}[\log \beta_j] \\
&- \sum_{i=1}^N \sum_{t=1}^T \frac{\partial}{\partial \beta_j}\mathbb{E}[V_{\mathbf{r}}(s_{i,t})] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \frac{\partial}{\partial \beta_j}\mathbb{E}[V_{\mathbf{r}}(s')].
\end{aligned}$$

Similarly to the previous section, we can handle all derivatives of expectations in the same way:

$$\begin{aligned}
\frac{\partial}{\partial \beta_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] &= \frac{\partial}{\partial \beta_j} \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) q(\boldsymbol{\lambda}) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u} \\
&= \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) q(\lambda_0) \cdots q(\lambda_{j-1}) \frac{\lambda_j^{\alpha_j - 1}}{\Gamma(\alpha_j)} \frac{\partial}{\partial \beta_j}[\beta_j^{\alpha_j} e^{-\beta_j \lambda_j}] \\
& \qquad q(\lambda_{j+1}) \cdots q(\lambda_d) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u}.
\end{aligned}$$

Since

$$\frac{\partial}{\partial \beta_j}[\beta_j^{\alpha_j} e^{-\beta_j \lambda_j}] = \alpha_j \beta_j^{\alpha_j - 1} e^{-\beta_j \lambda_j} - \beta_j^{\alpha_j} e^{-\beta_j \lambda_j} \lambda_j = \beta_j^{\alpha_j} e^{-\beta_j \lambda_j} \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right),$$

we have that

$$\begin{aligned}
\frac{\partial}{\partial \beta_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] = & \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) q(\lambda_0) \cdots q(\lambda_{j-1}) \frac{\beta_j^{\alpha_j}}{\Gamma(\alpha_j)} \lambda_j^{\alpha_j - 1} e^{-\beta_j \lambda_j} \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \\
& \qquad q(\lambda_{j+1}) \cdots q(\lambda_d) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u} \\
= & \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) q(\boldsymbol{\lambda}) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u} \\
= & \mathbb{E}\left[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] = \frac{\alpha_j}{\beta_j} \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] - \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \lambda_j].
\end{aligned}$$

This gives us the final expression of $\frac{\partial \mathcal{L}}{\partial \beta_j}$:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \beta_j} = & -\frac{1}{\beta_j} - \frac{1}{2}\mathbb{E}\left[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}] \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] - \frac{1}{2}\mathbb{E}\left[\mathrm{tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}] \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] \\
& - \frac{1}{2}\mathbb{E}\left[\mathbf{m}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] - \frac{1}{2}\mathbb{E}\left[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}| \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] \\
& + \mathbb{E}\left[\mathbf{t}^\intercal \mathbf{K}_{\mathbf{r},\mathbf{u}}^\intercal \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] \\
& - \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{E}\left[V_{\mathbf{r}}(s_{i,t}) \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \mathbb{E}\left[V_{\mathbf{r}}(s') \left(\frac{\alpha_j}{\beta_j} - \lambda_j \right) \right].
\end{aligned}$$

References

[1] H. Herrlich. *Axiom of choice*. Springer, 2006.

[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.

[4] H. Royden and P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010.