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

Paulius Dilkas (2146879)

16th 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* [4]:

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

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 [5], 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

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

$\square$

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.

Lemma 1.4 (Perturbation Lemma [3])**.** *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*

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

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.

Proposition 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

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

for all $s \in \mathcal{S}$. Let $s_0 \in \mathcal{S}$ be an arbitrary state. In order to prove that $V(s_0)$ is measurable, it is enough to show that for any $\alpha \in \mathbb{R}$, the set

$$\left\{ \mathbf{r} \in \mathbb{R}^{|\mathcal{S}|} \;\middle|\; \begin{aligned} & V_{\mathbf{r}}(s_0) \in (-\infty, \alpha); \\ & V_{\mathbf{r}}(s) \in \mathbb{R} \text{ for all } s \in \mathcal{S} \setminus \{s_0\}; \\ & V_{\mathbf{r}}(s) = \log \sum_{a \in \mathcal{A}} \exp\left(r(s) + \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s, a, s') V_{\mathbf{r}}(s')\right) \text{ for all } s \in \mathcal{S} \end{aligned} \right\}$$

is measurable. Since this set can be constructed in Zermelo-Fraenkel set theory *without* the axiom of choice, it is measurable [2], which proves that $V(s)$ is a measurable function for any $s \in \mathcal{S}$. $\square$

Proposition 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). ☐

Theorem 2.3 (The Lebesgue Dominated Convergence Theorem [6])**.** *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

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

Definition 2.4. Let $\mathbf{x} = (x_1, \ldots, x_n)^\intercal$ be a vector. For any coefficients $a_{i,j} \in \mathbb{R}$, we will call

$$\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j$$

a *quadratic form* of $\mathbf{x}$.

Lemma 2.5. *For $i = 0, \ldots, d$,*

$$q(\mathbf{r})(a + b(\mathbf{u})) \leq \left.\frac{\partial q(\mathbf{r})}{\partial \lambda_i}\right|_{\lambda_i = c(\mathbf{r},\mathbf{u})} \leq q(\mathbf{r})(c + d(\mathbf{u})),$$

where $a, c > 0$ are constants independent of $\mathbf{r}$ and $\mathbf{u}$, and $b(\mathbf{u})$, $d(\mathbf{u})$ are quadratic forms.

Proof. Since

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

by Lemma 1.1, we need to bound the elements of

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

If $i = 0$, then each element of $\frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_0}$ is of the form

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

i.e., without λ_0, so

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

is already independent of $\mathbf{r}$ and $\mathbf{u}$—there is no need for any bound.

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 $c(\mathbf{u}) > \lambda_i - \epsilon$,

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

which gives an upper bound on each element of

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

A similar line of reasoning establishes lower bounds as well. Combining this result with...

Since we already found an upper bound for

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

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

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

also exists. □

Remark. In order to find $\frac{\partial q(\mathbf{u})}{\partial t}$, where t is the ith element of the vector $\mathbf{m}$, we can find $\frac{\partial q(\mathbf{u})}{\partial \mathbf{m}}$ and simply take the ith element. A similar line of reasoning applies to matrices as well. Thus, we only need to consider derivatives with respect to $\mathbf{m}$, $\mathbf{S}$, and λ_i (for $i = 0, \ldots, d$).

Lemma 2.6. *Let t denote any scalar part of $\mathbf{m}$ or $\mathbf{S}$. Then*

$$\left| \left. \frac{\partial q(\mathbf{u})}{\partial t} \right|_{t = c(\mathbf{r},\mathbf{u})} \right| \leq c q(\mathbf{u}),$$

where $c > 0$ is independent of $\mathbf{r}$ and $\mathbf{u}$.

Proof. Since

$$\frac{\partial q(\mathbf{u})}{\partial \mathbf{m}} = q(\mathbf{u}) \frac{1}{2} \operatorname{tr} \left((\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{u} (\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{u})^\intercal - \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}) \left. \frac{\partial \mathbf{K}_{\mathbf{u},\mathbf{u}}}{\partial \lambda_i} \right|_{\lambda_i = c(\mathbf{u})} \right)$$

by Lemma 1.1.

Hereinafter we fix i and j to refer respectively to the row and column of t in case of derivative by a matrix, and we fix i to be the index of t in $\mathbf{m}$ if the derivative is taken with respect to a vector. We will analyse each of the three cases separately.

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).

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.

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 $k \in \{1, \ldots, m\} \setminus \{i\}$,

$$\mathbb{E}[\mathbf{u} - \mathbf{c}(\mathbf{u})]_k = \mathbb{E}[u_k - c_k] = m_k - c_k,$$

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).

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 $\mathbf{E}(\mathbf{u})$ is a matrix 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

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

1. Applying Lemma 1.4 to $\mathbf{S}$ and $\mathbf{E}(\mathbf{u})$ gives

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

   which can be reformulated to

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

   Choosing to use the maximum norm we get

   $$\max_k \sum_l \left|[\mathbf{C}(\mathbf{u})^{-1}]_{k,l} - [\mathbf{S}^{-1}]_{k,l}\right| \leq \|\mathbf{S}^{-1}\|_\infty^2 |c(\mathbf{u}) - t|.$$

   Using (5) we get that for any row k,

   $$\sum_l \left|[\mathbf{C}(\mathbf{u})^{-1}]_{k,l} - [\mathbf{S}^{-1}]_{k,l}\right| < \|\mathbf{S}^{-1}\|_\infty^2 \epsilon,$$

   and for any row k and column l,

   $$\left|[\mathbf{C}(\mathbf{u})^{-1}]_{k,l} - [\mathbf{S}^{-1}]_{k,l}\right| < \|\mathbf{S}^{-1}\|_\infty^2 \epsilon,$$

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

2. Because of the result in 1, 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 $\mathbb{E}[\mathbf{u}\mathbf{u}^\intercal]$.

$\square$

Our main theorem is a specialised version of an integral differentiation result by Chen [1].

Theorem 2.7. *Whenever the derivative exists,*

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

where t is any scalar part of $\mathbf{m}$, $\mathbf{S}$, 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 fix the value of 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 Proposition 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. For the time being, we can assume that t is a parameter of $q(\mathbf{r})$. Then

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

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

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

for some function $c : \mathbb{R}^{|\mathcal{S}|}\times\mathbb{R}^m \to (\min\{t,t_n\},\max\{t,t_n\})$ 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

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

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

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

We then have that

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

The bound is clearly non-negative and measurable. It remains to show that it is also integrable. Let $c > 0$ be as given by either Lemma 2.5 or Lemma 2.6. Then

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

Since $\|\mathbf{r}\|_\infty \leq \|\mathbf{r}\|_1$, and

$$\iint \|\mathbf{r}\|_1 q(\mathbf{r}) q(\mathbf{u})\, d\mathbf{r}\, d\mathbf{u} = \sum_{i=1}^{|\mathcal{S}|} \mathbb{E}[|r_i|]$$

results in a finite number, we get the required result. □

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

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

and

$$\begin{aligned} \frac{\partial}{\partial\mathbf{m}}\mathbb{E}[V_\mathbf{r}(s)] &= \frac{\partial}{\partial\mathbf{m}} \iiint V_\mathbf{r}(s) p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X}_\mathbf{u}, \mathbf{u}) \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) q(\boldsymbol{\lambda})\, d\mathbf{r}\, d\mathbf{u}\, d\boldsymbol{\lambda} \\ &= \iiint V_\mathbf{r}(s) p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X}_\mathbf{u}, \mathbf{u}) \frac{\partial}{\partial\mathbf{m}} \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) q(\boldsymbol{\lambda})\, d\mathbf{r}\, d\mathbf{u}\, d\boldsymbol{\lambda}, \end{aligned} \tag{7}$$

where Substituting it back into (7) 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* [4]. 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

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

3.3 $\partial/\partial\alpha_j$

We begin in the usual way:

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

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:

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

Then

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

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

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

3.4 $\partial/\partial\beta_j$

Finally,

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

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 $\frac{\partial \mathcal{L}}{\partial \beta_j}$:

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

References

[1] R. Chen. The dominated convergence theorem and applications. National Cheng Kung University, 2016.

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

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

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

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

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