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

Paulius Dilkas (2146879)

3rd December 2018

1 Proofs

Theorem 1. *MDP value functions $V(s) : \mathbb{R}^{|\mathcal{S}|} \to \mathbb{R}$ (for $s \in \mathcal{S}$) are Lebesgue measurable.*

Definition 1. *For any finite-dimensional vector $\mathbf{x} = (x_1, \dots, x_n)^\intercal$, its* maximum norm *is*

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

Theorem 2. *If the initial values of the MDP value function satisfy the following bound, then the bound remains satisfied throughout value iteration:*

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

Theorem 3 (The Lebesgue Dominated Convergence Theorem [2]). *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 1 ([2]). *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 4. *Using our usual notation,*

$$\frac{\partial}{\partial t} \iiint V_r(s) q(\mathbf{r}) q(\mathbf{u}) q(\boldsymbol{\lambda}) \, d\mathbf{r} \, d\mathbf{u} \, d\boldsymbol{\lambda} = \iiint \frac{\partial}{\partial t} [V_r(s) q(\mathbf{r}) q(\mathbf{u}) q(\boldsymbol{\lambda})] \, d\mathbf{r} \, d\mathbf{u} \, d\boldsymbol{\lambda},$$

where $t \in \{\mathbf{m}, \mathbf{S}, \alpha_0, \dots, \alpha_d, \beta_0, \dots, \beta_d\}$.

Proof. Let

$$f(\mathbf{r}, \mathbf{u}, \boldsymbol{\lambda}, t) = V_r(s) q(\mathbf{r}) q(\mathbf{u}) q(\boldsymbol{\lambda}),$$

$$F(t) = \iiint f(\mathbf{r}, \mathbf{u}, \boldsymbol{\lambda}, t) \, d\mathbf{r} \, d\mathbf{u} \, d\boldsymbol{\lambda},$$

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

$$F'(t) = \lim_{n\to\infty} \frac{F(t_n) - F(t)}{t_n - t} = \iiint \left. \frac{\partial f}{\partial t} \right|_{(\mathbf{r}, \mathbf{u}, \boldsymbol{\lambda}, t)} d\mathbf{r}\, d\mathbf{u}\, d\boldsymbol{\lambda}. \tag{1}$$

We have

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

where

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

Since

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

(1) follows from Theorem 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, $\mathbf{r}$, $\mathbf{u}$, $\boldsymbol{\lambda}$,

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

The MDP value function is measurable by Theorem 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 $q(\boldsymbol{\lambda})$. Then

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

since PDFs are non-negative. An upper bound for $|V_r(s)|$ is given by Theorem 2, while

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

for some c between t and 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). Therefore,

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

The bound is clearly non-negative and measurable. It remains to show that it is also integrable. Let $\mathbf{r} = (r_1, \ldots, r_k)^\intercal$. Then

$$\frac{\|\mathbf{r}\|_\infty + \log|\mathcal{A}|}{1 - \gamma} q(\mathbf{r}) \leq \frac{q(\mathbf{r})}{1 - \gamma} \left(\log|\mathcal{A}| + \sum_{i=1}^{k} |r_i| \right),$$

and

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

which clearly exists and is finite, so

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

for all γ and $\mathbf{u}$. The existence of

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

then comes from a boundedness argument, as... $\qquad \square$

Theorem 5 (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 $t_n \ne t$ for each t. We claim that

$$\lim_{n \to \infty} \frac{F(t_n) - F(t)}{t_n - t} = \int_R \frac{\partial f(x,t)}{\partial t} \, d\mu(x). \tag{2}$$

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 (2) 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. $\qquad\square$

2   Derivatives of the Evidence Lower Bound

2.1 $\partial/\partial\mathbf{m}$

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

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

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

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

(3)

where

$$\frac{\partial}{\partial \mathbf{m}} \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) = \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) \frac{\partial}{\partial \mathbf{m}} \left[-\frac{1}{2} (\mathbf{u} - \mathbf{m})^\mathsf{T} \mathbf{S}^{-1} (\mathbf{u} - \mathbf{m}) \right]$$
$$= \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) \left(-\frac{1}{2} \right) (\mathbf{S}^{-1} + \mathbf{S}^{-\mathsf{T}}) (\mathbf{u} - \mathbf{m}) \frac{\partial}{\partial \mathbf{m}} [\mathbf{u} - \mathbf{m}]$$
$$= \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) \frac{1}{2} (\mathbf{S}^{-1} + \mathbf{S}^{-\mathsf{T}}) (\mathbf{u} - \mathbf{m}).$$

Substituting it back into (3) gives

$$\frac{\partial}{\partial \mathbf{m}} \mathbb{E}[V_r(s)] = \frac{1}{2} \iiint V_r(s) (\mathbf{S}^{-1} + \mathbf{S}^{-\mathsf{T}}) (\mathbf{u} - \mathbf{m}) p(\mathbf{r}|\boldsymbol{\lambda}, \mathbf{X_u}, \mathbf{u}) \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) q(\boldsymbol{\lambda}) \, d\mathbf{r} \, d\mathbf{u} \, d\boldsymbol{\lambda}$$
$$= \frac{1}{2} \mathbb{E}[V_r(s) (\mathbf{S}^{-1} + \mathbf{S}^{-\mathsf{T}}) (\mathbf{u} - \mathbf{m})].$$

Hence

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

2.2 $\partial / \partial \mathbf{S}$

Similarly to the previous section,

$$\frac{\partial \mathcal{L}}{\partial \mathbf{S}} = \frac{1}{2} \frac{\partial}{\partial \mathbf{S}} \log |\mathbf{S}| - \frac{1}{2} \frac{\partial}{\partial \mathbf{S}} \mathrm{Tr}[\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] \mathbf{S}]$$
$$- \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{\partial}{\partial \mathbf{S}} \mathbb{E}[V_r(s_{i,t})] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \frac{\partial}{\partial \mathbf{S}} \mathbb{E}[V_r(s')],$$

where

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

and

$$\frac{\partial}{\partial \mathbf{S}} \mathrm{Tr}[\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}] \mathbf{S}] = \mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^\mathsf{T}$$

by *The Matrix Cookbook* [1]. Then

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

4

where

$$\frac{\partial}{\partial \mathbf{S}} \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) = \frac{\partial}{\partial \mathbf{S}} \left[\frac{1}{(2\pi)^{m/2}|\mathbf{S}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m}) \right) \right]$$

$$= \frac{\partial}{\partial \mathbf{S}} \left[\frac{1}{(2\pi)^{m/2}|\mathbf{S}|^{1/2}} \right] \exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m}) \right)$$

$$+ \frac{1}{(2\pi)^{m/2}|\mathbf{S}|^{1/2}} \frac{\partial}{\partial \mathbf{S}} \left[\exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m}) \right) \right]$$

$$= \frac{1}{(2\pi)^{m/2}} \frac{\partial}{\partial \mathbf{S}} \left[\frac{1}{|\mathbf{S}|^{1/2}} \right] \exp\left(-\frac{1}{2}(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m}) \right)$$

$$- \frac{1}{2}\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})\frac{\partial}{\partial \mathbf{S}}[(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m})].$$

The two remaining derivatives can be taken with the help of *The Matrix Cookbook* [1]:

$$\frac{\partial}{\partial \mathbf{S}} \left[\frac{1}{|\mathbf{S}|^{1/2}} \right] = -\frac{1}{2}|\mathbf{S}|^{-3/2}\frac{\partial |\mathbf{S}|}{\partial \mathbf{S}} = -\frac{1}{2}|\mathbf{S}|^{-3/2}|\mathbf{S}|\mathbf{S}^{-\mathsf{T}} = -\frac{1}{2|\mathbf{S}|^{1/2}}\mathbf{S}^{-\mathsf{T}},$$

$$\frac{\partial}{\partial \mathbf{S}}[(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-1}(\mathbf{u} - \mathbf{m})] = -\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}}.$$

Plugging them back in gives

$$\frac{\partial}{\partial \mathbf{S}} \mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) = -\frac{1}{2}\mathbf{S}^{-\mathsf{T}}\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S}) + \frac{1}{2}\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}},$$

and

$$\frac{\partial}{\partial \mathbf{S}} \mathbb{E}[V_r(s)] = \frac{1}{2} \iiint V_r(s)(\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}} - \mathbf{S}^{-\mathsf{T}})q(\mathbf{r})\mathcal{N}(\mathbf{u}; \mathbf{m}, \mathbf{S})q(\boldsymbol{\lambda}) \, d\mathbf{r} \, d\mathbf{u} \, d\boldsymbol{\lambda}$$

$$= \frac{1}{2}\mathbb{E}[V_r(s)(\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}} - \mathbf{S}^{-\mathsf{T}})].$$

Therefore

$$\frac{\partial \mathcal{L}}{\partial \mathbf{S}} = \frac{1}{2}\mathbf{S}^{-\mathsf{T}} - \frac{1}{2}\mathbb{E}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}]^{\mathsf{T}} - \frac{1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}[V_r(s_{i,t})(\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}} - \mathbf{S}^{-\mathsf{T}})]$$

$$- \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s')\mathbb{E}[V_r(s')(\mathbf{S}^{-\mathsf{T}}(\mathbf{u} - \mathbf{m})(\mathbf{u} - \mathbf{m})^{\mathsf{T}}\mathbf{S}^{-\mathsf{T}} - \mathbf{S}^{-\mathsf{T}})].$$

2.3 $\partial/\partial \alpha_j$

We begin in the usual way:

$$\frac{\partial \mathcal{L}}{\partial \alpha_j} = -\frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\text{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\text{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathbf{m}^{\mathsf{T}}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}] - \frac{1}{2}\frac{\partial}{\partial \alpha_j}\mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}|]$$

$$+ \frac{\partial}{\partial \alpha_j}\mathbb{E}[\mathbf{t}^{\mathsf{T}}\mathbf{K}_{\mathbf{r},\mathbf{u}}^{\mathsf{T}}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}] + \frac{\partial}{\partial \alpha_j}[\alpha_j + \log \Gamma(\alpha_j) + (1 - \alpha_j)\psi(\alpha_j)]$$

$$- \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\partial}{\partial \alpha_j}\mathbb{E}[V_r(s_{i,t})] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s')\frac{\partial}{\partial \alpha_j}\mathbb{E}[V_r(s')].$$

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

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

$$
q(\lambda_{j+1}) \cdots q(\lambda_d) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u}
$$

$$
= \iiint f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\log \beta_j + \log \lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) q(\boldsymbol{\lambda}) q(\mathbf{r}) q(\mathbf{u}) \, d\boldsymbol{\lambda} \, d\mathbf{r} \, d\mathbf{u}
$$

$$
= \mathbb{E}\left[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \left(\log \beta_j + \log \lambda_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \right]
$$

$$
= \left(\log \beta_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r})] + \mathbb{E}[f(k_{\boldsymbol{\lambda}}, \mathbf{r}) \log \lambda_j].
$$

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

$$
\frac{\partial \mathcal{L}}{\partial \alpha_j} = 1 + (1 - \alpha_j) \psi'(\alpha_j) + \left(\log \beta_j - \frac{\Gamma'(\alpha_j)}{\Gamma(\alpha_j)} \right) \left(-\frac{1}{2} \mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2} \mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{S}]] \right.
$$

$$
- \frac{1}{2} \mathbb{E}[\mathbf{m}^{\mathsf{T}} \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}] - \frac{1}{2} \mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}|] + \mathbb{E}[\mathbf{t}^{\mathsf{T}} \mathbf{K}_{\mathbf{r},\mathbf{u}}^{\mathsf{T}} \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m}]
$$

$$
\left. - \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{E}[V_r(s_{i,t})] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \mathbb{E}[V_r(s')] \right)
$$

$$
- \frac{1}{2} \mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}] \log \lambda_j] - \frac{1}{2} \mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{S}] \log \lambda_j] - \frac{1}{2} \mathbb{E}[\mathbf{m}^{\mathsf{T}} \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \log \lambda_j]
$$

$$
- \frac{1}{2} \mathbb{E}[\log |\mathbf{K}_{\mathbf{u},\mathbf{u}}| \log \lambda_j] + \mathbb{E}[\mathbf{t}^{\mathsf{T}} \mathbf{K}_{\mathbf{r},\mathbf{u}}^{\mathsf{T}} \mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1} \mathbf{m} \log \lambda_j]
$$

$$
- \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{E}[V_r(s_{i,t}) \log \lambda_j] - \gamma \sum_{s' \in \mathcal{S}} \mathcal{T}(s_{i,t}, a_{i,t}, s') \mathbb{E}[V_r(s') \log \lambda_j].
$$

2.4 $\partial/\partial\beta_j$

Finally,

$$\frac{\partial\mathcal{L}}{\partial\beta_j} = -\frac{1}{2}\frac{\partial}{\partial\beta_j}\mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]] - \frac{1}{2}\frac{\partial}{\partial\beta_j}\mathbb{E}[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]] - \frac{1}{2}\frac{\partial}{\partial\beta_j}\mathbb{E}[\mathbf{m}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}]$$

$$-\frac{1}{2}\frac{\partial}{\partial\beta_j}\mathbb{E}[\log|\mathbf{K}_{\mathbf{u},\mathbf{u}}|] + \frac{\partial}{\partial\beta_j}\mathbb{E}[\mathbf{t}^{\intercal}\mathbf{K}_{\mathbf{r},\mathbf{u}}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}] - \frac{\partial}{\partial\beta_j}[\log\beta_j]$$

$$-\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\partial}{\partial\beta_j}\mathbb{E}[V_r(s_{i,t})] - \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t},a_{i,t},s')\frac{\partial}{\partial\beta_j}\mathbb{E}[V_r(s')].$$

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

$$\frac{\partial}{\partial\beta_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}},\mathbf{r})] = \frac{\partial}{\partial\beta_j}\iiint f(k_{\boldsymbol{\lambda}},\mathbf{r})q(\boldsymbol{\lambda})q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}$$

$$= \iiint f(k_{\boldsymbol{\lambda}},\mathbf{r})q(\lambda_0)\cdots q(\lambda_{j-1})\frac{\lambda_j^{\alpha_j-1}}{\Gamma(\alpha_j)}\frac{\partial}{\partial\beta_j}[\beta_j^{\alpha_j}e^{-\beta_j\lambda_j}]$$

$$q(\lambda_{j+1})\cdots q(\lambda_d)q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}.$$

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

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

$$q(\lambda_{j+1})\cdots q(\lambda_d)q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}$$

$$= \iiint f(k_{\boldsymbol{\lambda}},\mathbf{r})\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)q(\boldsymbol{\lambda})q(\mathbf{r})q(\mathbf{u})\,d\boldsymbol{\lambda}\,d\mathbf{r}\,d\mathbf{u}$$

$$= \mathbb{E}\left[f(k_{\boldsymbol{\lambda}},\mathbf{r})\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right] = \frac{\alpha_j}{\beta_j}\mathbb{E}[f(k_{\boldsymbol{\lambda}},\mathbf{r})] - \mathbb{E}[f(k_{\boldsymbol{\lambda}},\mathbf{r})\lambda_j].$$

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

$$\frac{\partial\mathcal{L}}{\partial\beta_j} = -\frac{1}{\beta_j} - \frac{1}{2}\mathbb{E}\left[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-2}]\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right] - \frac{1}{2}\mathbb{E}\left[\mathrm{Tr}[\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{S}]\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right]$$

$$-\frac{1}{2}\mathbb{E}\left[\mathbf{m}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right] - \frac{1}{2}\mathbb{E}\left[\log|\mathbf{K}_{\mathbf{u},\mathbf{u}}|\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right]$$

$$+\mathbb{E}\left[\mathbf{t}^{\intercal}\mathbf{K}_{\mathbf{r},\mathbf{u}}^{\intercal}\mathbf{K}_{\mathbf{u},\mathbf{u}}^{-1}\mathbf{m}\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right]$$

$$-\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\left[V_r(s_{i,t})\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right] - \gamma\sum_{s'\in\mathcal{S}}\mathcal{T}(s_{i,t},a_{i,t},s')\mathbb{E}\left[V_r(s')\left(\frac{\alpha_j}{\beta_j} - \lambda_j\right)\right].$$

References

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

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