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

Paulius Dilkas (2146879)

16th November 2018

1 Derivatives of the Evidence Lower Bound

1.1 $\partial/\partial\mathbf{m}$

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

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

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

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

where

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

Substituting it back into (1) gives

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

Hence

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

1.2 $\partial/\partial\mathbf{S}$

Similarly to the previous section,

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

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* [1]. Then

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

where

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

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

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

Plugging them back in gives

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

and

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

Therefore

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

1.3 $\partial/\partial\alpha_j$

We begin in the usual way:

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

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:

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

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

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

1.4 $\partial/\partial\beta_j$

Finally,

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

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

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

2 Proofs

Theorem 1 (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. Applying the mean value theorem to the function $t\mapsto f(x,t)$, for each $t=t_0$ we have to have some c between t_0 and t so that

$$f(x,t)-f(x,t_0) = \left.\frac{\partial f}{\partial t}\right|_{(x,c)}(t-t_0).$$

It follows that

$$|f(x,t)-f(x,t_0)|\le g(x)|t-t_0|$$

and so

$$|f(x,t)|\le|f(x,t_0)|+g(x)|t-t_0|.$$

Thus

$$\int_R|f(x,t)|\,d\mu(x)\le\int(|f(x,t_0)|+g(x)|t-t_0|)\,d\mu(x) = \int_R|f(x,t_0)|\,d\mu(x)+|t-t_0|\int_R g\,d\mu<\infty,$$

which establishes that the function $x \mapsto f(x,t)$ is integrable for each t.

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

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

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

References

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