Department of Mathematics
Technical University of Munich
Elisabeth Ullmann

October 10, 2024

Solution Sheet 1 – Importance sampling

Autumn School *Uncertainty Quantification for High-dimensional Problems*

Problem 1 (Warm up)

a) Let f be the pdf of the uniform random variable U on the interval $[0,1]$. Let g be the pdf of the uniform random variable V on the interval $[0,1/2]$. Let $H(u)=u^2$. Show that $\mathbb{E}[H(U)] \neq \mathbb{E}[H(V)f(V)/g(V)]$.

b) Let f be the pdf of the univariate standard normal density and $H(u)=\exp(-(u-10)^2/2)$. Find the optimal importance sampling density!

c) Let f be the pdf of the univariate standard normal density and $H(u)=\exp(ku)$, where $k \neq 0$. Find the optimal importance sampling density!

Solution

a) $\mathbb{E}[H(U)] = \int_0^1 u^2 \mathrm{d}u = 1/3$ and $\mathbb{E}[H(V)f(V)/g(V)] = \frac{2}{2}\int_0^{1/2} u^2 \mathrm{d}u = 1/24$.

b) The optimal IS density is (up to a normalizing constant)

$$H(u)f(u) = \exp(-(u-10)^2/2)\exp(-u^2/2) = \exp(-(u^2-10u+50)) \propto \exp(-(u-5)^2).$$

This is the normal density with mean $\mu=5$ and variance $\sigma^2=1/2$.

c) The optimal IS density is (up to a normalizing constant)

$$H(u)f(u) = \exp(ku)\exp(-u^2/2) = \exp(-(u^2-2ku+k^2-k^2)/2) \propto \exp(-(u-k)^2/2).$$

This is the normal density with mean $\mu=k$ and variance $\sigma^2=1$.

Problem 2 (Self-normalized importance sampling)

Let $\boldsymbol{U} \sim f$ be a random vector with pdf f. Consider estimating

$$Q := \mathbb{E}[H(\boldsymbol{U})] = \int_{D_f} H(\boldsymbol{u}) f(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}$$

with the self-normalized importance sampling estimator

$$E^{IS}_{sn,g}[Q] = \frac{\frac{1}{N}\sum_{i=1}^{N} W(\boldsymbol{V}^{(i)}) H(\boldsymbol{V}^{(i)})}{\frac{1}{N}\sum_{i=1}^{N} W(\boldsymbol{V}^{(i)})},$$

where $W(\boldsymbol{u}) = f(\boldsymbol{u})/g(\boldsymbol{u})$ is the likelihood ratio and $\boldsymbol{V}^{(i)} \sim g$ i.i.d. for $i = 1, \ldots, N$. In this problem we assume that the density g dominates the density f, that is, $\mathcal{D}_f \subseteq \mathcal{D}_g$.

a) Let $\boldsymbol{Z}^{(i)}$ be i.i.d. copies of a random vector taking values in $\mathbb{R}^n$ with distribution $\mathbb{P}_{\boldsymbol{Z}}$. Let $\overline{\boldsymbol{Z}} := \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{Z}^{(i)}$ denote the Monte Carlo estimator of $\mathbb{E}[\boldsymbol{Z}]$. Let $v \colon \mathbb{R}^n \to \mathbb{R}$ denote a smooth function. The *delta method* approximates $v(\overline{\boldsymbol{Z}})$ by a truncated Taylor expansion of v with anchor point $v(\mathbb{E}[\boldsymbol{Z}])$ as follows:

$$\widetilde{v}(\overline{\boldsymbol{Z}}) := v(\mathbb{E}[\boldsymbol{Z}]) + \nabla v(\mathbb{E}[\boldsymbol{Z}])^\top (\overline{\boldsymbol{Z}} - \mathbb{E}[\boldsymbol{Z}]).$$

Show that the variance of $\widetilde{v}(\overline{\boldsymbol{Z}})$ is given by

$$\mathrm{var}(\widetilde{v}(\overline{\boldsymbol{Z}})) = \frac{1}{N} \nabla v(\mathbb{E}[\boldsymbol{Z}])^\top \operatorname{Cov}(\boldsymbol{Z}, \boldsymbol{Z}) \nabla v(\mathbb{E}[\boldsymbol{Z}]).$$

b) Let $\sigma^2_{sn,g} := \mathbb{E}[W(\boldsymbol{V})^2 (H(\boldsymbol{V}) - Q)^2]$, where $\boldsymbol{V} \sim g$ is a random vector with pdf g. Show that the delta method approximates the variance of the self-normalized IS estimator as follows

$$\mathrm{var}(E^{IS}_{sn,g}[Q]) \approx \frac{\sigma^2_{sn,g}}{N}. \tag{1}$$

c) Show that the importance sampling density which minimizes the approximate variance of $E^{IS}_{sn,g}[Q]$ in (1) is given by

$$g_{opt,sn}(\boldsymbol{u}) = \frac{|H(\boldsymbol{u}) - Q| f(\boldsymbol{u})}{\int |H(\boldsymbol{u}) - Q| f(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}}.$$

d) Show that

$$\sigma^2_{sn,g} \geq \mathbb{E}[|H(\boldsymbol{U}) - Q|]^2.$$

e) Finally, let $H(\boldsymbol{u}) = \mathbb{1}_{\{G \leq 0\}}(\boldsymbol{u})$ be the indicator function of a failure domain with probability of failure $P_f = Q$. Which lower bound for $\sigma^2_{sn,g}$ do we obtain in this case? Derive an (approximate) lower bound for the c.o.v. of the self-normalized IS estimator! Compare this bound with the c.o.v. of the standard Monte Carlo estimator for P_f!

Solution

a) Note that by construction

$$\mathbb{E}\left[\widetilde{v}(\overline{\boldsymbol{Z}})\right] = v(\mathbb{E}\left[\boldsymbol{Z}\right]) + \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top (\underbrace{\mathbb{E}\left[\overline{\boldsymbol{Z}}\right]}_{=\mathbb{E}[\boldsymbol{Z}]} - \mathbb{E}\left[\boldsymbol{Z}\right]) = v(\mathbb{E}\left[\boldsymbol{Z}\right]) + \mathbf{0} = v(\mathbb{E}\left[\boldsymbol{Z}\right]).$$

Hence we obtain

$$\begin{aligned}
\operatorname{var}(\widetilde{v}(\overline{\boldsymbol{Z}})) &= \mathbb{E}\left[(v(\mathbb{E}\left[\boldsymbol{Z}\right]) + \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right]) - \mathbb{E}\left[\widetilde{v}(\overline{\boldsymbol{Z}})\right])^2\right] \\
&= \mathbb{E}\left[(\nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right]))^2\right] \\
&= \mathbb{E}\left[\nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right])(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right])^\top \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])\right] \\
&= \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top \underbrace{\mathbb{E}\left[(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right])(\overline{\boldsymbol{Z}} - \mathbb{E}\left[\boldsymbol{Z}\right])^\top\right]}_{=\operatorname{Cov}(\boldsymbol{Z},\boldsymbol{Z})/N} \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right]) \\
&= \frac{1}{N}\nabla v(\mathbb{E}\left[\boldsymbol{Z}\right])^\top \operatorname{Cov}\left(\boldsymbol{Z},\boldsymbol{Z}\right) \nabla v(\mathbb{E}\left[\boldsymbol{Z}\right]).
\end{aligned}$$

b) For the self-normalized IS estimator we have $n = 2$, $v(z_1, z_2) = z_2/z_1$, $\overline{Z}_1 = \frac{1}{N}\sum_{i=1}^N W(\boldsymbol{V}^{(i)})$ and $\overline{Z}_2 = \frac{1}{N}\sum_{i=1}^N H(\boldsymbol{V}^{(i)})W(\boldsymbol{V}^{(i)})$, where $\boldsymbol{V}, \boldsymbol{V}^{(i)} \sim g$ i.i.d. Moreover, $Z_1 = W(\boldsymbol{V}^{(1)})$, $\mathbb{E}\left[Z_1\right] = \mathbb{E}[W(\boldsymbol{V}^{(1)})] = 1$, $Z_2 = H(\boldsymbol{V}^{(1)})W(\boldsymbol{V}^{(1)})$ and $\mathbb{E}\left[Z_2\right] = \mathbb{E}[H(\boldsymbol{V}^{(1)})W(\boldsymbol{V}^{(1)})] = \mathbb{E}[H(\boldsymbol{U})] = Q$. The gradient

$$\nabla v(\boldsymbol{z}) = (-z_2/z_1^2, 1/z_1)^\top$$

and

$$\nabla v(\mathbb{E}\left[\boldsymbol{Z}\right]) = (-\mathbb{E}[H(\boldsymbol{V})W(\boldsymbol{V})]/\mathbb{E}[W(\boldsymbol{V})]^2, 1/\mathbb{E}[W(\boldsymbol{V})])^\top = (-Q, 1)^\top.$$

Hence we obtain the (approximate) variance

$$\begin{aligned}
\operatorname{var}(\widetilde{v}(\overline{\boldsymbol{Z}})) &= \frac{1}{N}(-Q, 1)^\top \operatorname{Cov}(\boldsymbol{Z}, \boldsymbol{Z})(-Q, 1) \\
&= \frac{1}{N}(Q^2 \operatorname{var}(Z_1) - 2Q\operatorname{Cov}(Z_1, Z_2) + \operatorname{var}(Z_2)) \\
&= \frac{1}{N}\operatorname{var}[(Z_2 - QZ_1)^2] \\
&= \frac{1}{N}\mathbb{E}[(Z_2 - QZ_1)^2] - \frac{1}{N}\underbrace{\mathbb{E}[(Z_2 - QZ_1)]^2}_{=0} \\
&= \frac{1}{N}\mathbb{E}[(H(\boldsymbol{V})W(\boldsymbol{V}) - QW(\boldsymbol{V}))^2] = \frac{1}{N}\mathbb{E}[W(\boldsymbol{V})^2(H(\boldsymbol{V}) - Q)^2].
\end{aligned}$$

c) The proof is analogous to the proof for the (ordinary) importance sampling estimator $E_g^{IS}[Q]$, where we minimize $\mathbb{E}[W(\boldsymbol{V})^2 H(\boldsymbol{V})^2]$ with respect to the density g. For the self-normalized IS estimator we minimize $\mathbb{E}[W(\boldsymbol{V})^2(H(\boldsymbol{V}) - Q)^2]$ with respect to g.

d) By choosing $g = g_{opt}$ derived in c) we obtain a lower bound for $\sigma^2_{sn,g}$ as follows:

$$\begin{aligned}
\sigma^2_{sn,g} &= \mathbb{E}[W^2(\boldsymbol{V})(H(\boldsymbol{V}) - Q)^2] \\
&\geq \mathbb{E}[f^2(\boldsymbol{V}_{opt})(H(\boldsymbol{V}_{opt}) - Q)^2 / g_{opt}(\boldsymbol{V}_{opt})^2] \\
&= \int_{D_{g_{opt}}} \frac{f(\boldsymbol{u})^2 (H(\boldsymbol{u}) - Q)^2}{|H(\boldsymbol{u}) - Q| f(\boldsymbol{u})} \mathrm{d}\boldsymbol{u} \left(\int_{D_{g_{opt}}} |H(\boldsymbol{u}) - Q| f(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \right) \\
&= \left(\int_{D_{g_{opt}}} |H(\boldsymbol{u}) - Q| f(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} \right)^2 \\
&= \left(\int_{D_f} |H(\boldsymbol{u}) - Q| f(\boldsymbol{u}) \mathrm{d}\boldsymbol{u} + \int_{D_f^c \cap D_{g_{opt}}} |H(\boldsymbol{u}) - Q| \underbrace{f(\boldsymbol{u})}_{=0} \mathrm{d}\boldsymbol{u} \right)^2 \\
&= \mathbb{E}[|H(\boldsymbol{U}) - Q|]^2.
\end{aligned}$$

e) It holds $\mathbb{E}[|H(\boldsymbol{U}) - Q|]^2 = (|1 - P_f| P_f + |0 - P_f|(1 - P_f))^2 = 4P_f^2(1 - P_f)^2$. Assuming that $\mathbb{E}[E^{IS}_{sn,g}[P_f]] = P_f$, this gives a lower bound for the approximate c.o.v. of $E^{IS}_{sn,g}[P_f]$ as follows:

$$\mathrm{c.o.v.}(E^{IS}_{sn,g}[P_f]) \geq \sqrt{\frac{4P_f^2(1 - P_f)^2}{P_f^2 N}} = \frac{2(1 - P_f)}{\sqrt{N}}.$$

For the standard Monte Carlo estimator we have

$$\mathrm{c.o.v.}(E^{MC}[P_f]) = \sqrt{\frac{(1 - P_f)}{P_f N}}.$$