Importance Sampling

Yutao Chen

⧖ Aug 26 2025
⧗ Oct 01 2025

Contents

Say we are interested in evaluating the following expectation:

$$\mathbb{E}_{p(\boldsymbol{x})}[f(\boldsymbol{x})] = \int p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x}.$$

A naive Monte Carlo approximation

$$\mathbb{E}_{p(\boldsymbol{x})}[f(\boldsymbol{x})] \approx \frac{1}{N} \sum_{n=1}^{N} p(\boldsymbol{x}_n) f(\boldsymbol{x}_n)$$

typically requires us being able to sample $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ from the *target distribution* $p(\boldsymbol{x})$.

Direct Importance Sampling

However, it is generally difficult to sample from arbitrary $p(\boldsymbol{x})$. Instead we resort to some *proposal distribution* $q(\boldsymbol{x})$:

$$\begin{aligned} \mathbb{E}_{p(\boldsymbol{x})}[f(\boldsymbol{x})] &= \int p(\boldsymbol{x}) f(\boldsymbol{x}) d\boldsymbol{x} \\ &= \int q(\boldsymbol{x}) \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} f(\boldsymbol{x}) d\boldsymbol{x} \\ &= \mathbb{E}_{q(\boldsymbol{x})}\left[\frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} f(\boldsymbol{x})\right] \\ &\approx \frac{1}{N} \sum_{n=1}^{N} w_n f(\boldsymbol{x}_n), \end{aligned}$$

where $w_n = p(\boldsymbol{x}_n) / q(\boldsymbol{x}_n)$ is referred to as the *importance weight*.

In this way, we bypass sampling from $p(\boldsymbol{x})$ by instead sampling from the proposal distribution $q(\boldsymbol{x})$. However, we still need to evaluate the densities $p(\boldsymbol{x})$ or $q(\boldsymbol{x})$.

Self-Normalized Importance Sampling

In some cases, even evaluating the densities $p(\boldsymbol{x})$ or $q(\boldsymbol{x})$ is difficult. In this case, we instead consider the unnormalized densities $\tilde{p}(\boldsymbol{x})$ and $\tilde{q}(\boldsymbol{x})$ such that

$$
\begin{aligned}
p(\boldsymbol{x}) &= \tilde{p}(\boldsymbol{x})/Z_p, \\
q(\boldsymbol{x}) &= \tilde{q}(\boldsymbol{x})/Z_q,
\end{aligned}
$$

where $Z_p = \int \tilde{p}(\boldsymbol{x})d\boldsymbol{x}$ and $Z_q = \int \tilde{q}(\boldsymbol{x})d\boldsymbol{x}$. We now have

$$
\begin{aligned}
\mathbb{E}_{p(\boldsymbol{x})}[f(\boldsymbol{x})] &= \int p(\boldsymbol{x})f(\boldsymbol{x})d\boldsymbol{x} \\
&= \frac{Z_q}{Z_p}\int q(\boldsymbol{x})\frac{\tilde{p}(\boldsymbol{x})}{\tilde{q}(\boldsymbol{x})}f(\boldsymbol{x})d\boldsymbol{x} \\
&= \frac{Z_q}{Z_p}\mathbb{E}_{q(\boldsymbol{x})}\left[\frac{\tilde{p}(\boldsymbol{x})}{\tilde{q}(\boldsymbol{x})}f(\boldsymbol{x})\right] \\
&\approx \frac{Z_q}{Z_p}\frac{1}{N}\sum_{n=1}^{N}\tilde{w}_n f(\boldsymbol{x}_n),
\end{aligned}
$$

where $\tilde{w}_n = \tilde{p}(\boldsymbol{x}_n)/\tilde{q}(\boldsymbol{x}_n)$. Similarly, we can derive that

$$
\frac{Z_p}{Z_q} = \int q(\boldsymbol{x})\frac{\tilde{p}(\boldsymbol{x})}{\tilde{q}(\boldsymbol{x})}d\boldsymbol{x} \approx \frac{1}{N}\sum_{n=1}^{N}\tilde{w}_n.
$$

Putting everything together, we have the following simplified form of *self-normalized importance sampling*:

$$
\mathbb{E}_{p(x)}[f(x)] \approx \sum_{n=1}^{N}\left(\frac{\tilde{w}_n}{\sum_{m=1}^{N}\tilde{w}_m}f(\boldsymbol{x}_n)\right).
$$