Monte Carlo Integration

Monte Carlo integration is a numerical integration method which derives from the idea of expectation in probability theory.

Monte Carlo Method

Idea:
The expectation of a function of a random variable is

$$\mathbb{E}[f(x)] = \int_D f(x)p(x)\,dx \simeq \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

which gives the idea of Monte Carlo integration

$$\int_D f(x)\,dx = \int_D \frac{f(x)}{p(x)}p(x)\,dx \simeq \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)} = F_N$$

Example:

$$F = \int_a^b e^{\sin(3x^2)}\,dx \simeq F_N = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}$$

where we use uniform distribution, i.e. $p(x_i) = \frac{1}{b-a}$. Below is the code for integration

```C++
double integrate(int N, double a, double b)
{
        double x, sum=0.0;
        for (int i = 0; i < N; ++i)
        {
                x = a + drand48()*(b-a);
                sum += exp(sin(3*x*x));
        }

        return sum / double(N);
}
```

Monte Carlo Estimator
The statistic F_N is called Monte Carlo estimator

- unbiased estimator
- extension to higher dimensions straightforward
- convergence in $O\left(\frac{1}{\sqrt{N}}\right)$ cause $V[F_N] = \frac{1}{N}V[F_1]$ ⇒ variance ~ $1/N$, std ~ $1/\sqrt{N}$

Proof of Convergence Rate

$$
\begin{aligned}
V\left[F_N\right] &= V\left[\frac{1}{N} \sum_{i=0}^{N} \frac{f\left(X_i\right)}{p\left(X_i\right)}\right] = \frac{1}{N^2} V\left[\sum_{i=0}^{N} \frac{f\left(X_i\right)}{p\left(X_i\right)}\right] \\
&= \frac{1}{N^2} \sum_{i=0}^{N} V\left[\frac{f\left(X_i\right)}{p\left(X_i\right)}\right] = \frac{1}{N} V\left[\frac{f(X)}{p(X)}\right] = \frac{1}{N} V\left[F_1\right]
\end{aligned}
$$

Pros:

- flexible
- easy to implement
- easily handle complex integrands
- efficient for high dimensional integrands

Cons:

- variance(noise)
- slow convergence $O\left(\frac{1}{\sqrt{N}}\right)$

Importance Sampling

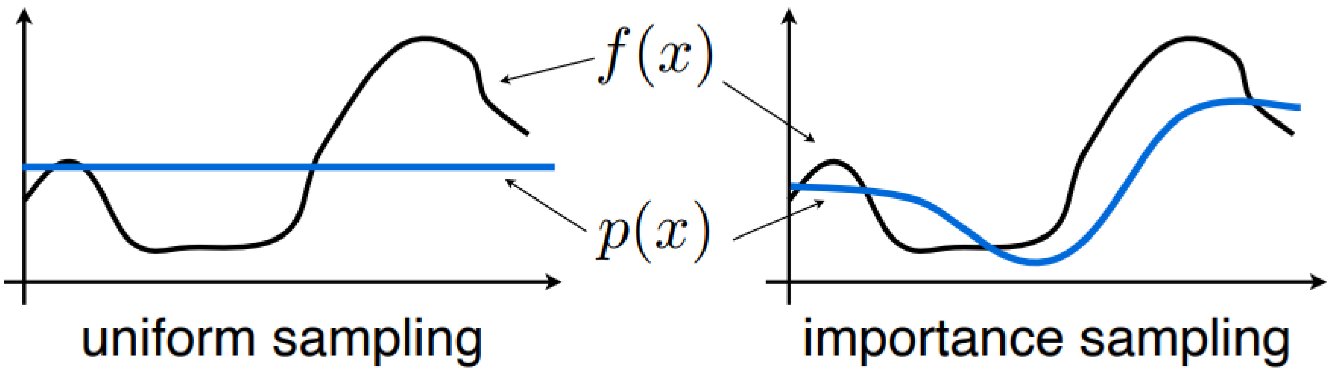
Assume $p(x) = cf(x)$

$$
\int p(x)\, dx = 1 \rightarrow c = \frac{1}{\int f(x)\, dx}
$$

Monte Carlo estimator is

$$
\frac{f(X_i)}{p(X_i)} = \frac{1}{c} = \int f(x)\, dx
$$

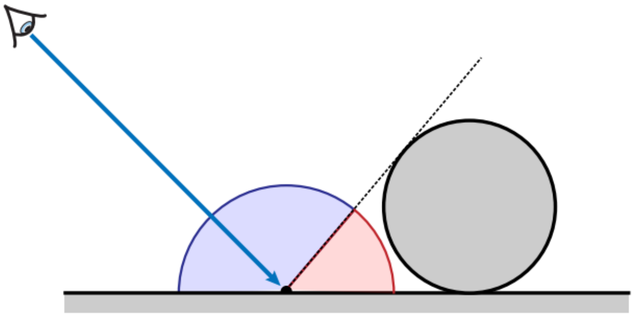
which is infeasible since we don't know the integral. But if pdf is similar to integrand, variance can be significantly reduced



Example: Ambient Occlusion

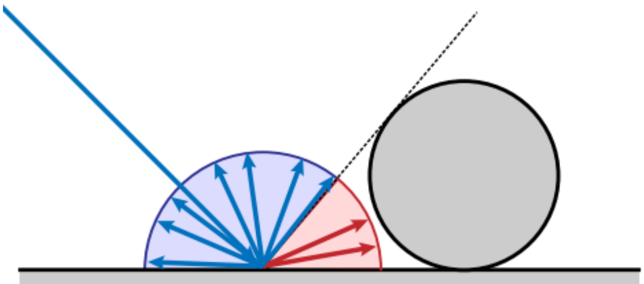
Consider diffuse objects illuminated by an ambient white sky, the rendering function is as below and can be simplified by the fact that the BRDF and incident light are constant.

$$
\begin{aligned}
L_r\left(\mathbf{x}, \vec{\omega}_r\right) &= \int_{H^2} f_r\left(\mathbf{x}, \vec{\omega}_i, \vec{\omega}_r\right) L_i\left(\mathbf{x}, \vec{\omega}_i\right) \cos\theta_i d\vec{\omega}_i \\
L_r(\mathbf{x}) &= \frac{\rho}{\pi} \int_{H^2} V\left(\mathbf{x}, \vec{\omega}_i\right) \cos\theta_i d\vec{\omega}_i
\end{aligned}
$$



$$L_r(\mathbf{x}) \approx \frac{\rho}{\pi N} \sum_{k=1}^{N} \frac{V\left(\mathbf{x}, \vec{\omega}_{i,k}\right) \cos\theta_{i,k}}{p\left(\vec{\omega}_{i,k}\right)}$$

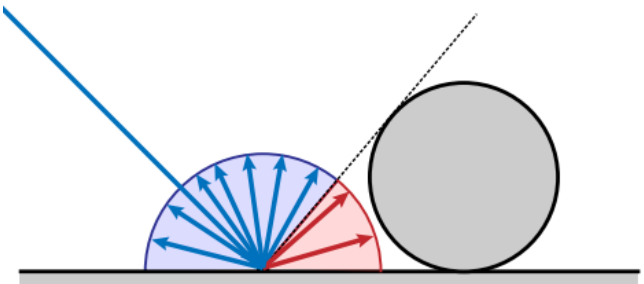
Uniform hemispherical sampling



$$p\left(\vec{\omega}_{i,k}\right) = 1/2\pi$$

$$L_r(\mathbf{x}) \approx \frac{2\rho}{N} \sum_{k=1}^{N} V\left(\mathbf{x}, \vec{\omega}_{i,k}\right) \cos\theta_{i,k}$$

Cosine-weighted importance sampling



$$p\left(\vec{\omega}_{i,k}\right) = \cos\theta_{i,k}/\pi$$

$$L_r(\mathbf{x}) \approx \frac{\rho}{N} \sum_{k=1}^{N} V\left(\mathbf{x}, \vec{\omega}_{i,k}\right)$$

Combining Multiple Strategies

We could sample from the average PDF

$$\frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{0.5(p_1(x_i) + p_2(x_i))}$$

Note that it's no use to average two different estimators:

$$\frac{0.5}{N_1} \sum_{i=1}^{N_1} \frac{f\left(x_i\right)}{p_1\left(x_i\right)} + \frac{0.5}{N_2} \sum_{i=1}^{N_2} \frac{f\left(x_i\right)}{p_2\left(x_i\right)}$$

since variance is additive.

Multiple Importance Sampling(MIS)

In MC integration, variance is high when the PDF is not proportional to the integrand

Ways of Combination

Naïve combination of 2 sampling strategies(no use):
Sample N_1 points using PDF p_1 then sample N_2 points using PDF p_2

$$\langle F^{N_1+N_2} \rangle = \frac{w_1}{N_1} \sum_{i=1}^{N_1} \frac{f(x_i)}{p_1(x_i)} + \frac{w_2}{N_2} \sum_{i=1}^{N_2} \frac{f(x_i)}{p_2(x_i)}$$

Weighted combination of 2 sampling strategies

$$\langle F^{N_1+N_2} \rangle = \frac{1}{N_1} \sum_{i=1}^{N_1} w_1(x_i) \frac{f(x_i)}{p_1(x_i)} + \frac{1}{N_2} \sum_{i=1}^{N_2} w_2(x_i) \frac{f(x_i)}{p_2(x_i)}$$

where $w_1(x) + w_2(x) = 1$. Note that the weight here is different for every point.

Weighted combination of M sampling strategies

$$\left\langle F^{\sum N_s} \right\rangle = \sum_{s=1}^{M} \frac{1}{N_s} \sum_{i=1}^{N_s} w_s(x_i) \frac{f(x_i)}{p_s(x_i)}$$

where $\sum_{s=1}^{M} w_s(x) = 1$

Choice of the Weights

- Balance heuristic(provably good)

$$w_s(x) = \frac{N_s p_s(x)}{\sum_j N_j p_j(x)}$$

- Power heuristic

$$w_s(x) = \frac{(N_s p_s(x))^\beta}{\sum_j (N_j p_j(x))^\beta}$$

One-Sample Model

$$\langle F^1 \rangle = w_s(x) \frac{f(x)}{q_s p_s(x)}$$

where q_s is the probability of using s-th strategy. If we derive q_s from the multi-sample model we have

$$q_s = \frac{N_s}{\sum N_j}$$

And the balance heuristic for the one-sample mode gives

$$w_s(x) = \frac{N_s p_s(x)}{\sum_j N_j p_j(x)} = \frac{q_s p_s(x)}{\sum_j q_j p_j(x)}$$

Plug into the one-sample model we have

$$\langle F^1 \rangle = w_s(x)\frac{f(x)}{q_s p_s(x)} = \frac{q_s p_s(x)}{\sum_j q_j p_j(x)}\frac{f(x)}{q_s p_s(x)} = \frac{f(x)}{\sum_j q_j p_j(x)}$$

which turns out that the probability is a linear combination of PDFs.

In a word, if we want to use multiple sample strategy, define a new PDF to be their linear combination.

$$\langle F^N \rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{\sum_j q_j p_j(x_i)}$$

The strategy works because if we have a large value $f(x_i)$ we should also have a relatively large value in the denominator(as long as at least one PDF is large)