SUNG-EUI YOON, KAIST

RENDERING

FREELY AVAILABLE ON THE INTERNET

FREELY AVAILABLE ON THE INTERNET

http://sglab.kaist.ac.kr/~sungeui/render

First printing, July 2018

14
Monte Carlo Integration

In this chapter, we study Monte Carlo integration to evaluate complex integral functions such as our rendering equation. In the next chapter, we will discuss Monte Carlo based ray tracing techniques that are specialized techniques for evaluating the rendering equations.

The rendering equation (Eq. 13.1) is a complex integration function. First of all, to compute a radiance for a ray starting from a surface point x, we need to integrate all the incoming radiances that arrive at x. Moreover, evaluating those incoming radiances requires us to evaluate the same procedure in a recursive way. Since there could be an infinite number of light paths starting from a light source to the eye, it is almost impossible to find an analytic solution for the rendering equation, except simple cases.

Rendering equations can be high dimensional, since we need to consider motion blur and many other effects with time and complex camera lens.

Second, the rendering equation can be high dimensional. The rendering equation shown in Eq. 13.1 is two dimensional. In practice, we need to support the motion blur for dynamic models and moving cameras. Considering such motion blur, we need to integrate radiance over time in each pixel, resulting in three dimensional rendering equation. Furthermore, supporting realistic cameras requires two or more additional dimensions on the equation. As a result, the equation for generating realistic images and video could be five or more dimensional.

Due to these issues, high dimensionality and infinite number of possible light paths, deriving analytic solutions and using deterministic approaches such as quadrature rules are impossible for virtually all of rendering environments that we encounter. Monte Carlo integration was proposed to integrate such high-dimensional functions based on random samples.

Overall, Monte Carlo (MC) integration is a numerical solution to integrate high complex and high-dimensional function. Since it uses sampling, it has stochastic errors, commonly quantified as Mean Squared Error (MSE). Fortunately, MC integration is unbiased, indicating that it gives us a correct solution with an infinite number of samples on average.

14.1 MC Estimator

Suppose that we have the following integration, whose solution is I:

$$I = \int_a^b f(x)dx. \tag{14.1}$$

The goal of MC integration is to take N different random samples, x_i, that follow the same probability density function, $p(x_i)$. We then use the following estimator:

$$\hat{I} = \frac{1}{N}\sum_i \frac{f(x_i)}{p(x_i)}. \tag{14.2}$$

We now discuss how the MC estimator is good. One of measures for this goal is Mean Squared Error (MSE), measuring the difference between the estimated values, $\hat{Y}_i$, and observed, real values, Y_i:

$$MSE(\hat{Y}) = E[(\hat{Y} - Y)^2] = \frac{1}{N}\sum_i (\hat{Y}_i - Y_i)^2. \tag{14.3}$$

MSE can be decomposed into bias and variances terms as the following:

$$MSE(\hat{Y}) = E\left[\left(\hat{Y} - E[\hat{Y}]\right)^2\right] + \left(E(\hat{Y}) - Y\right)^2 \tag{14.4}$$

$$= Var(\hat{Y}) + Bias(\hat{Y}, Y)^2. \tag{14.5}$$

The bias term $Bias(\hat{Y}, Y)$ measures how much the average value of the estimator $\hat{Y}$ is away from its ground-truth value Y. On other hand, the variance term $Var(\hat{Y})$ measures how the estimator values are away from its average values. We would like to discuss bias and variance of the MC estimator (Eq. 14.2).

Bias of the MC estimator. The MC estimator is unbiased, i.e., on average, it returns the correct solution, as shown in below:

$$\begin{aligned}
E[\hat{I}] &= E\left[\frac{1}{N}\sum_i \frac{f(x_i)}{p(x_i)}\right] \\
&= \frac{1}{N}\int \sum_i \frac{f(x_i)}{p(x_i)} p(x)dx \\
&= \frac{1}{N}\sum_i \int \frac{f(x)}{p(x)} p(x)dx, \because x_i \text{ samples have the same } p(x) \\
&= \frac{N}{N}\int f(x)dx = I.
\end{aligned} \tag{14.6}$$

Variance of the MC estimator. To derive the variance of the MC estimator, we utilize a few properties of variance. Based on those properties, and Independent and Identically Distributed samples (IID) of random samples, the variance of the MC estimator can be derived as the following:

$$\begin{aligned}
Var(\hat{I}) &= Var(\frac{1}{N}\sum_i \frac{f(x_i)}{p(x_i)}) \\
&= \frac{1}{N^2} Var(\sum_i \frac{f(x_i)}{p(x_i)}) \\
&= \frac{1}{N^2}\sum_i Var(\frac{f(x_i)}{p(x_i)}), \because x_i \text{ samples are independent from each other.} \\
&= \frac{1}{N^2} N Var(\frac{f(x)}{p(x)}), \because x_i \text{ samples are from the same distribution.} \\
&= \frac{1}{N} Var(\frac{f(x)}{p(x)}) = \frac{1}{N}\int \left(\frac{f(x)}{p(x)} - E\left[\frac{f(x)}{p(x)}\right]\right)^2 p(x)dx. \quad (14.7)
\end{aligned}$$

As can be in the above equations, the variance of the MC estimator decreases as a function of $\frac{1}{N}$, where N is the number of samples.

Simple experiments with MC estimators. Suppose that we would like to compute the following, simple integration:

$$I = \int_0^1 4x^3 dx = 1. \quad (14.8)$$

We know its ground truth value, 1, for the integration. We can now study various properties of the MC estimator by comparing its result against the ground truth. When we use the uniform sampling on the integration domain, the MC estimator is defined as the following:

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N} 4x_i^3, \quad (14.9)$$

where $p(x_i) = p_x = 1$, since the sampling domain is $[0, 1]$, and the integration of uniform sampling on the domain has to be one, $\int_0^1 p_x = 1$.

Fig. 14.1 shows how the MC estimator behaves as we have more samples, N. As can be seen, MC estimators approach to its ground truth value, as we have more samples. Furthermore, when we measure the mean and variance of different MC estimators that have different random numbers given the same MC estimator equation (Eq. 14.9), their mean and variance shows the expected behaviors; its mean is same to the ground truth and the variance decreases as a function of $\frac{1}{N}$, respectively.

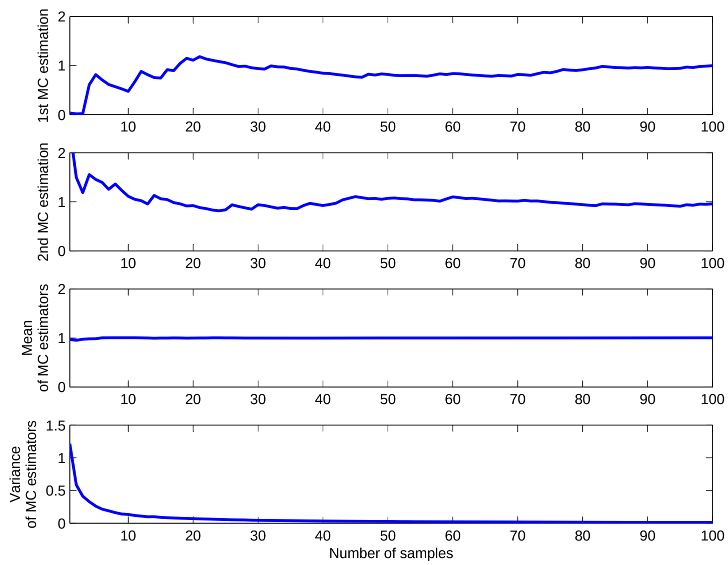

Figure 14.1: The top two subfigures show the first and second MC estimators of $\int_0^1 4x^3 dx$, whose ground truth value is 1. These MC estimators approach to their ground-truth, as we have more number of samples. While these individual MC estimators have up and down depending on their randomly generated values, their mean and variance measured with 600 estimators show the expected behavior, as theoretically predicted in Sec. 14.1. Its source code, mc_int_ex.m, is available.

14.2 *High Dimensions*

Suppose that we have an integration with higher dimensions than one:

$$I = \int \int f(x, y) dx dy. \tag{14.10}$$

Even in this case, our MC estimator is extended straightforwardly to handle such an two-dimensional integration (and other higher ones):

$$\hat{I} = \frac{1}{N} \sum \frac{f(x_i, y_i)}{p(x_i, y_i)}, \tag{14.11}$$

where we generate N random samples following a two dimensional probability density function, $p(x, y)$. We see how to generate samples according to pdf in Sec. 14.4. This demonstrates that MC integration supports well high dimensional integrations including the rendering equation with many integration domains, e.g., image positions, time, and lens parameters.

In addition, MC integration has the following characteristics:

- **Simplicity.** We can compute MC estimators based only on point sampling. This results in very convenient and simple computation.

- **Generality.** As long as we can compute values at particular points of functions under the integration, we can use MC estimations. As a result, we can compute integrations of discontinuous functions, high dimensional functions, etc.

Example. Suppose that we would like to compute the following integration defined over a hemisphere:

$$I = \int_{\Theta} f(\Theta) dw_{\Theta}, \tag{14.12}$$

$$= \int_0^{2\pi} \int_0^{\frac{\pi}{2}} f(\theta, \phi) \sin\theta d\theta d\phi. \tag{14.13}$$

where Θ is the hemispherical coordinates, (θ, ϕ).

The MC estimator for the above integration can be defined as follows:

$$\hat{I} = \frac{1}{N} \sum \frac{f(\theta_i, \phi_i) \sin\theta}{p(\theta_i, \phi_i)}, \tag{14.14}$$

where we generate (θ_i, ϕ_i) following $p(\theta_i, \phi_i)$.

Now let's get back to the irradiance example mentioned in Sec. 12.2. The irradiance equation we discussed in the irradiance example is to use $L_s \cos\theta$ for $f(\theta, \phi)$. In this case, the MC estimator of Eq. 14.14 is transformed to:

$$\hat{I} = \frac{1}{N} \sum \frac{L_s \cos\theta \sin\theta}{p(\theta_i, \phi_i)}. \tag{14.15}$$

One can use different pdf $p(\theta, \phi)$ for the MC estimator, but we can use the following one:

$$p(\theta_i, \phi_i) = \frac{\cos\theta \sin\theta}{\pi}, \tag{14.16}$$

where the integration of the pdf in the domain is one: i.e., $\int_0^{2\pi} \int_0^{\frac{\pi}{2}} \cos\theta \sin\theta = 1$. Plugging the pdf into the estimator of Eq. 14.14, we get the following:

$$\hat{I} = \frac{\pi}{N} \sum L_s. \tag{14.17}$$

14.3 *Importance Sampling*

In this section, we see how different pdfs affect variance of our MC estimators. As we see in Sec. 14.1, our MC estimator is unbiased regardless of pdf employed, i.e., its mean value becomes the ground truth of the integration. Variances, however, vary depending on chosen pdf.

Let's see the example integration, $I = \int_0^1 4x^3 dx = 1$, again. In the following, we test three different pdfs and see their variance:

- $p(x) = 1$. As the simplest choice, we can use the uniform distribution on the domain. The variance of our MC estimator, $\hat{I} = \frac{1}{N} \sum_i 4x_i^3$ is $\frac{36}{28N} \approx \frac{1.285}{N}$, according to the variance equation (Eq. 14.7).

- $p(x) = x$. The variance of this MC estimator, $\frac{1}{N}\sum_i 4x^2$, is $\frac{14}{12N} \approx \frac{1.666}{N}$. Its variance is reduced from the above, uniform pdf!
- $p(x) = 4x^3$. The shape of this pdf is same to the underlying function under the integration. In this case, its variance turns out to be zero.

As demonstrated in the above examples, the variance of a pdf decreases, as the distribution of a pdf gets closer to the underlying function $f(x)$. Actually, when the pdf $p(x)$ is set to be $\frac{f(x)}{\int f(x)dx} = \frac{f(x)}{I}$, the ideal distribution, we get the lowest variance, zero. This can be shown as the following:

The variance of an MC estimator goes to zero, when the shape of its pdf is same to the underlying function under the integration. We, however, do not know such a shape of the rendering equation!

$$\begin{aligned} Var(\hat{I}) &= \frac{1}{N}\int \left(\frac{f(x)}{p(x)} - I\right)^2 p(x)dx \\ &= \frac{1}{N}\int (I - I)^2 p(x)dx \\ &= 0. \end{aligned} \tag{14.18}$$

Unfortunately, in some cases, we do not know the shape of the function under the integration. Especially, this is the case for the rendering equation. Nonetheless, the general idea is to generate more samples on high values on the function, since this can reduce the variance of our MC estimator, as demonstrated in aforementioned examples. In the same reason, when the pdf is chosen badly, the variance of our MC estimator can even go higher.

This is the main idea of importance sampling, i.e., generate more samples on high values on the underlying function, resulting in a lower variance.

The main idea of importance sampling is to generate more samples on high values on the function.

Fortunately, we can intuitively know which regions we can get high values on the rendering equation. For example, for the light sources, we can get high radiance values, and we need to generate rays toward such light sources to reduce the variance in our MC estimators. Technical details on importance sampling are available in Ch. 14.3.

14.4 Generating Samples

We can use any pdf for the MC estimator. In the case of the uniform distribution, we can use a random number generator, which generates random numbers uniformly given a range.

The question that we would like to ask in this section is how we can generate samples according to the pdf $p(x)$ different from the uniform pdf.

Fig. 14.2 shows a pdf and its cdf (cumulative distribution function) in a discrete setting. Suppose that we would like to generate samples

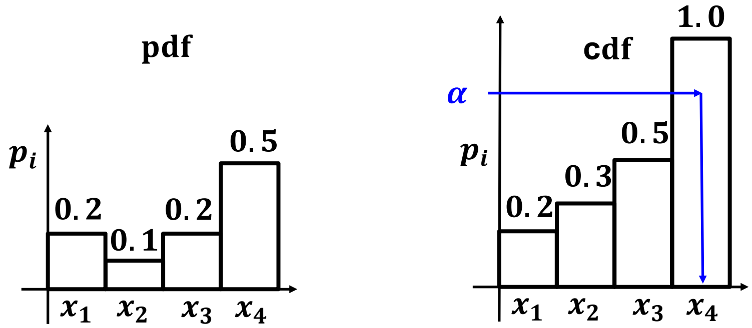

Figure 14.2: This figure shows a pdf and its cdf. Using the inverse cumulative distribution function generates samples according to the pdf by utilizing its cdf.

according to the pdf. In this case, x_1, x_2, x_3, x_4 are four events, whose probabilities are $0.2, 0.1, 0.2, 0.5$, respectively. In other words, we would like to generate those events with the pre-defined pdf.

We can use an inverse cumulative distribution function to generate samples according to a pdf.

A simple method of generating samples according to the pdf is to utilize its cdf (Fig. 14.2). This is known to be inverse cumulative distribution function. In this method, we first generate a random number α uniformly in the rnage of [0, 1). When the random number α is in the range $[\sum_0^{i-1} p_i, \sum_0^i p_i)$, we return a sample of x_i.

Let's see the probability of generating a sample x_i in this way to be p_i, as the following:

$$\begin{aligned} p(x_i) &= p(\alpha \in [\sum_0^{i-1} p_i, \sum_0^{i} p_i]) \\ &= p(\sum_0^{i} p_i) - p(\sum_0^{i-1} p_i) \\ &= p_i, \end{aligned} \tag{14.19}$$

where p_0 is set to be zero. So far, we see the discrete case, and we now extend it to the continuous case.

Continuous case. Suppose that we have a pdf, $p(x)$. Its cdf function, $F_X(x)$, is defined as $F_X(x) = p(X < x) = \int_{-\infty}^{x} p(x)dx$. We then generate a random number α uniformly in a range $[0, 1]$. A sample, y, is generated as $y = F_X^{-1}(\alpha)$.

Example for the diffuse emitter. Let's consider the following integration of measuring the irradiance with the diffuse emitter and our sampling pdf:

$$\begin{aligned} I &= \frac{1}{\pi}\int_{\Theta} dw_{\Theta}, \\ &= \frac{1}{\pi}\int_0^{2\pi}\int_0^{\frac{\pi}{2}} \sin\theta\cos\theta d\theta d\phi. \end{aligned} \tag{14.20}$$

$$p(\theta,\phi) = \frac{\sin\theta\cos\theta}{\pi}, \tag{14.21}$$

where $\int\int p(\theta,\phi)d\theta d\phi = 1$.

Our goal is to generate samples according to the chosen pdf. We first compute its cdf, $CDF(\theta,\phi)$, as the following:

$$\begin{aligned} CDF(\theta,\phi) &= \int_0^{\phi}\int_0^{\theta} \frac{\sin\theta\cos\theta}{\pi} d\theta d\phi \\ &= (1-\cos^2\theta)\frac{\phi}{2\pi} = F(\theta)F(\pi), \end{aligned} \tag{14.22}$$

where $F(\theta)$ and $F(\pi)$ are $(1-\cos^2\theta)$ and $\frac{\phi}{2\pi}$, respectively. Since the pdf is two dimensional, we generate two random numbers, α and β. We then utilize inverse function of those two separated functions of $F(\theta)$ and $F(\phi)$:

$$\begin{aligned} \theta &= F^{-1}(\alpha) = \cos^{-1}\sqrt{1-\alpha}, \\ \phi &= F^{-1}(\beta) = 2\pi\beta. \end{aligned} \tag{14.23}$$

The aforementioned, the inverse CDF method assumes that we can compute the inverse of the CDF. In some cases, we cannot compute the inverse of CDFs, and thus cannot use the inverse CDF method. In this case, we can use the rejection method.

In the rejection method, we first generate two random numbers, α and β. We accept β, only when $\alpha \leq p(\beta)$ (Fig. 14.3). In the example of Fig. 14.3, the ranges of α and β are $[0,1]$ and $[a,b]$. In this approach, we can generate random numbers β according to the pdf $p(x)$ without using its cdf. Nonetheless, this approach can be inefficient, especially when we do not accept and thus reject samples. This inefficiency occurs when the value of $p(x)$ is smaller than the upper bound, which we generate such random numbers up to. The upper bound of α in our example is 1.

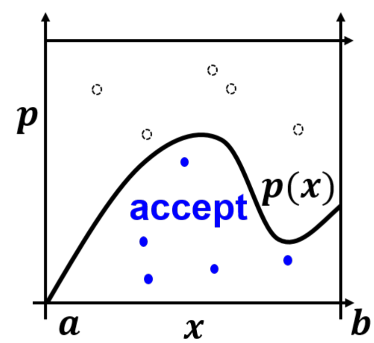

Figure 14.3: In the rejection method, we generate random numbers and accept numbers only when those numbers are within the pdf $p(x)$.