Monte Carlo Integration II
& Sampling from PDFs

CS295, Spring 2017

Shuang Zhao
Computer Science Department
University of California, Irvine
Last Lecture

• Direct illumination

• Monte Carlo integration I
Today’s Lecture

• Monte Carlo integration II
  • Convergence properties
  • Integrals over higher-dimensional domains

• Sampling from PDFs
  • Inversion method
Recap: Monte Carlo Integration

• Goal: to estimate $I = \int_{a}^{b} f(x) \, dx$

  • Pick a probability density function $p(x)$

  • Generate $n$ independent samples:

    $x_1, x_2, \ldots, x_n \sim p$

  • Evaluate $\hat{I}_j := \frac{f(x_j)}{p(x_j)}$ for $j = 1, 2, \ldots, n$

  • Return sample mean: $\bar{I} := \frac{1}{n} \sum_{j=1}^{n} \hat{I}_j$

How fast does this algorithm converge?
Central Limit Theorem

- Let \( \{X_1, X_2, \ldots\} \) be a sequence of i.i.d. random variables with \( \mathbb{E}[X_i] = \mu, \ Var[X_i] = \sigma^2 < \infty \) for all \( i \).

- Let \( S_n := \frac{1}{n} \sum_{i=1}^{n} X_i \), then \( \frac{\sqrt{n}}{\sigma}(S_n - \mu) \xrightarrow{d} N(0, 1) \).

- It follows that for any \( z > 0 \):
  \[
  \lim_{n \to \infty} \mathbb{P} \left[ \frac{\sqrt{n}}{\sigma}(S_n - \mu) \leq z \right] = \Phi(z)
  \]

CDF of standard normal distrb.
Central Limit Theorem

\[
\lim_{n \to \infty} \mathbb{P} \left[ \frac{\sqrt{n}}{\sigma} (S_n - \mu) \leq z \right] = \Phi(z)
\]

\[
\lim_{n \to \infty} \mathbb{P} \left[ \left| \frac{\sqrt{n}}{\sigma} (S_n - \mu) \right| \leq z \right] = \lim_{n \to \infty} \mathbb{P} \left[ \mu \in \left( S_n - z \frac{\sigma}{\sqrt{n}}, S_n + z \frac{\sigma}{\sqrt{n}} \right) \right] = 1 - 2\Phi(-z)
\]
Confidence Interval

- \( \left( S_n - z \frac{\sigma}{\sqrt{n}}, \ S_n + z \frac{\sigma}{\sqrt{n}} \right) \) is called a \( 1 - 2\Phi(-z) \) confidence interval (CI) of \( \mu \)

- **Interpretation:** there is a \( 1 - 2\Phi(-z) \) chance for \( \mu \) to reside in this CI

- \( z \) is called the “critical value”

- **Common critical values**
  - \( z = 1.960 \): 95%-CIs
  - \( z = 2.576 \): 99%-CIs
Confidence Interval

- \( \left( S_n - z \frac{\sigma}{\sqrt{n}}, \ S_n + z \frac{\sigma}{\sqrt{n}} \right) \) is called a \( 1 - 2\Phi(-z) \) confidence interval (CI) of \( \mu \)

- **One more problem:** the standard deviation \( \sigma \) is generally *unknown*

- **Solution:** use (corrected) sample standard deviation

\[
\sigma_n := \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - S_n)^2},
\]

resulting in CIs of the form \( \left( S_n - z \frac{\sigma_n}{\sqrt{n}}, \ S_n + z \frac{\sigma_n}{\sqrt{n}} \right) \)
Convergence of MC Methods

• The relative size of a CI, namely

\[
\left( z \frac{\sigma_n}{\sqrt{n}} \right) / |S_n|
\]

is an important indicator for the reliability of estimated result \( S_n \)

• Monte Carlo methods generally have a convergence rate of \( O(1/\sqrt{n}) \)
Today’s Lecture

• Monte Carlo integration II
  • Convergence properties
    • Integrals over higher-dimensional domains

• Sampling from PDFs
Background: Measures and Lebesgue Integration

\[ I = \int_{\Gamma} f(x) \, d\mu(x) \]

• \( \Gamma \) is a set specifying the domain of integration

• \( \mu \) is a measure function capturing the “sizes” of \( \Gamma \) and its (measurable) subsets
  • Properties:
    \[
    \mu(A) \geq 0 \quad \text{for all } A \subseteq \Gamma \text{ (that are measurable)}
    \]
    \[
    \mu(A \cup B) = \mu(A) + \mu(B) \quad \text{if } A \cap B = \emptyset
    \]
    \[
    \mu(\emptyset) = 0
    \]
Background: Measures and Lebesgue Integration

\[ I = \int_{\Gamma} f(x) \, d\mu(x) \]

- Assuming \( \Gamma \) can be partitioned into \( n \) disjoint components \( \Gamma_1, \Gamma_2, \ldots, \Gamma_n \) and \( f(x_i) \equiv a_i \) for all \( x_i \) in \( \Gamma_i \)

- Then, \( I = \sum_{i=1}^{n} a_i \mu(\Gamma_i) \)

- Special case: \( \int_{\Gamma} d\mu(x) = \mu(\Gamma) \)
Background: Measures and Lebesgue Integration

\[ I = \int_{\Gamma} f(x) \, d\mu(x) \]

- In general, \( \Gamma \) can be partitioned into many components \( \Gamma_1, \Gamma_2, \ldots \) so that \( f \) becomes (approximately) constant in each component \( \Gamma_i \).

- In the limiting case,

\[ I = \lim_{n \to \infty} \sum_{i=1}^{n} a_i \, \mu(\Gamma_i) \]
Background: Borel Measure and 1D Integrals

\[ \mu_B((a, b]) = b - a \]

- The *Borel measure* captures the “length” of 1D intervals

\[ \int_a^b f(x) \, dx = \int_{(a, b]} f(x) \, d\mu_B(x) \]
Background: 
Lebesgue Measure

\[ \lambda((a, b] \times (c, d]) = (b - a)(d - c) \]

• The Lebesgue measure captures the “volume” of hyper-cubes

\[ \int_{a}^{b} \int_{c}^{d} f((x, y)) \, dy \, dx = \int_{(a,b] \times (c,d]} f(r) \, d\lambda(r) \]
Background: Solid Angle Measure

- For a region $A$ on the unit sphere $\mathbb{S}^2$, the solid angle measure $\sigma$ captures $A$’s surface area.

Unit sphere

\[
\sigma(\mathbb{S}^2) = 4\pi
\]

Unit hemisphere

\[
\sigma(\Omega_+) = \sigma(\mathbb{H}^2) = 2\pi
\]
Background: Projected Solid Angle Measure

For a region $A$ on the unit sphere $S^2$, the projected solid angle measure $\sigma_\perp$ captures the surface area of $A_\perp$ (which is $A$’s projection on the plane perpendicular to the normal direction)

$$\sigma_\perp(\Omega_+) = \sigma_\perp(H^2) = \pi$$

$A$’s projection $A_\perp$ on a plane perpendicular to the normal direction $n$

$$d\sigma_\perp(\omega) = \cos \theta \, d\sigma(\omega)$$
Background: Probability Measures

- Given a sample space (i.e., all possible outcomes) \( \Omega \) and \( \mathcal{F} \subseteq \Omega \), \( P(\mathcal{F}) \) denotes the probability for an outcome in \( \mathcal{F} \) to occur

- Example 1-1: fair coin
  - \( \Omega = \{H, T\} \)
  - \( P(\{H\}) = P(\{T\}) = \frac{1}{2}, \quad P(\{H, T\}) = 1 \)

- Example 1-2: Poisson distribution (with parameter \( \lambda \))
  - \( \Omega = \{0, 1, 2, \ldots\} = \mathbb{N} \cup \{0\} \)
  - \( P(\{k\}) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \forall k \in \Omega \)
Background: Probability Measures

- Given a sample space (i.e., all possible outcomes) \( \Omega \) and \( \mathcal{F} \subseteq \Omega \), \( P(\mathcal{F}) \) denotes the probability for an outcome in \( \mathcal{F} \) to occur.

- Example 2-1: \( U(0, 1] \)
  - \( \Omega = (0, 1] \)
  - \( P((a, b]) = b - a = \mu_B((a, b]) \quad \forall 0 \leq a < b \leq 1 \)

- Example 2-2: \( N(0, 1) \)
  - \( \Omega = \mathbb{R} \)
  - \( P((a, b]) = \Phi(b) - \Phi(a) \quad \forall a < b \)
Background: Probability Measures

• Given a sample space (i.e., all possible outcomes) \( \Omega \) and \( \mathcal{F} \subseteq \Omega \), \( P(\mathcal{F}) \) denotes the probability for an outcome in \( \mathcal{F} \) to occur.

• Example 2-3: 1D continuous distrib. with CDF \( F \)
  - \( \Omega = \mathbb{R} \)
  - \( P((a, b]) = F(b) - F(a) \quad \forall a < b \)
  - \( P((-, a]) = F(a) \quad \forall a \in \mathbb{R} \)
  - There is a **one-to-one** correspondence between CDFs and probability measures.
Background: High-Dimensional Distributions

• Consider a sample space $\Omega \subseteq \mathbb{R}^k$ associated with a measure $\mu$

• E.g., $\Omega = \mathbb{R}^k$, $\mu = \lambda$ (Lebesgue);
  $\Omega = \mathbb{S}^2$, $\mu = \sigma$ (solid angle)

• Given a probability density function $p$ satisfying
  
  $p(x) \geq 0$ and $\int_{\Omega} p(x) \, d\mu(x) = 1$,

  the corresponding probability measure $P$ is

  $P(A) = \int_A p(x) \, d\mu(x)$

  with $dP(x) = p(x) \, d\mu(x)$
Background:
General Definition of Expected Values

\[ \mathbb{E}[f(x)] := \int_{\Omega} f(x) \, dP(x) \]

• Applies to both discrete and continuous cases

• Example 1: fair coin (discrete)
  • Consider \( f(H) = 2, \ f(T) = -1 \)
    (i.e., you win 2 dollars for an H and lose 1 for a T)

\[
\mathbb{E}[f(x)] = \int_{\{H,T\}} f(x) \, dP(x) \\
= f(H)P(\{H\}) + f(T)P(\{T\}) \\
= 2 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = \frac{1}{2}
\]
Background:
General Definition of Expected Values

\[
\mathbb{E}[f(x)] := \int_{\Omega} f(x) \, dP(x)
\]

• Example 2: \text{U}(0, 1] (continuous)
  • Consider \( f(x) = x \) for all \( x \) in \((0, 1]\)

\[
\mathbb{E}[f(x)] = \int_{(0,1]} f(x) \, dP(x) = \int_{(0,1]} x \, d\mu_B(x)
\]

\[
= \int_{0}^{1} x \, dx = \frac{1}{2}
\]
Background:
General Definition of Expected Values

\[ \mathbb{E}[f(\mathbf{x})] := \int_{\Omega} f(\mathbf{x}) \, dP(\mathbf{x}) \]

- General continuous distribution with:
  - Sample space \( \Omega \) and associated measure \( \mu \)
  - Probability density \( p \)

"Classical" definition of expected value

\[ \mathbb{E}[f(\mathbf{x})] = \int_{\Omega} f(\mathbf{x}) \, dP(\mathbf{x}) = \int_{\Omega} f(\mathbf{x}) p(\mathbf{x}) \, d\mu(\mathbf{x}) \]
Monte Carlo Integration over High-Dimensional Domains

\[ I = \int_{\Gamma} f(x) \, d\mu(x) \]

• Let \( p \) be a pdf on \( \Gamma \) (with \( f(x) > 0 \) implying \( p(x) > 0 \))

• Then

\[ \mathbb{E} \left[ \frac{f(x)}{p(x)} \right] = \int_{\Gamma} \frac{f(x)}{p(x)} \, p(x) \, d\mu(x) = I \]

• Exactly the same as the 1D case!
Today’s Lecture

• Monte Carlo integration II
  • Convergence properties
  • Integrals over higher-dimensional domains

• Sampling from PDFs
Sampling from PDFs

- Monte Carlo integration requires drawing independent samples from arbitrary probability density $p(x)$

- What we generally have in practice:
  - Generating random floating numbers from $U(0, 1]$
Inversion Method

• Given a 1D distribution with CDF \( F \), then

\[ X = F^{-1}(\xi) \quad \text{where} \quad \xi \sim U(0, 1) \]

follows this given distribution

• Proof: for all \( a \in \mathbb{R} \),

\[
\mathbb{P}[X < a] = \mathbb{P}[F^{-1}(\xi) < a]
= \mathbb{P}[\xi < F(a)]
= F(a)
\]
Inversion Method

• Also works for discrete distributions with finite outcomes: consider one with CMF $F$

• Sampling algorithm
  • Draw $\xi \sim U(0, 1]$  
  • Return the minimal integer $X$ satisfying $F(X) \geq \xi$
  • Complexity: $O(\log N)$
Next Lecture

• Sampling from PDFs II