Lecture 26, Mar 16, 2023

Monte Carlo Integration

- What if we want to numerically integrate a function in a large number of dimensions? $\int_{x_1} \cdots \int_{x_d} f(x_1, \cdots, x_d)\,\mathrm{d}x_d \cdots \mathrm{d}x_1$
  - In 1 dimension we can use the trapezoidal rule
  - The number of points we need to take scales exponentially with the number of dimensions, so for higher dimensions the number of points we need becomes impractical
  - The error of the trapezoidal approximation scales as $O(N^{-2})$ where N is the number of points per dimension; given N total points and d dimensions, the error would be $O((N^{1/d})^{-2}) = O(N^{-2/d})$
  - As the number of dimensions increase, the computational effort increases exponentially – "the curse of dimensionality"
- Consider if we had a sample of N uniform random variables in $[0,1]^d$, i.e. a d-dimensional unit cube, each random variable being a d-dimensional vector
  - This is a sample $\begin{bmatrix} x_1^1 \\ \vdots \\ x_d^1 \end{bmatrix}, \begin{bmatrix} x_1^2 \\ \vdots \\ x_d^2 \end{bmatrix}, \cdots, \begin{bmatrix} x_1^N \\ \vdots \\ x_d^N \end{bmatrix}$
- Let RV $Y = f(X_1, \cdots, X_d)$, where f is the function we want to integrate
  - $E[Y] = \int_0^1 \cdots \int_0^1 f(x_1, \cdots, x_d)\,\mathrm{d}x_1 \cdots \mathrm{d}x_d$
    - Since the distribution is uniform the distribution is 1 for all points in the unit cube
    - The expectation turns out to be the same as the integral we wanted to evaluate
- Let $\bar{Y} = \frac{1}{N}\sum_{k=1}^{N} Y^k = \frac{1}{N}\sum_{k=1}^{N} f(x_1^k, \cdots, x_d^k)$; this is our sample mean
  - $E[\bar{Y}] = \frac{1}{N}\sum_{k=1}^{N} E[f(x_1^k, \cdots, x_d^k)] = \frac{1}{N}\sum_{k=1}^{N}\mu = \mu$
    - This means $\bar{Y}$ is an unbiased estimator of $E[Y]$, which has the same value of the integral we want to solve
- Therefore we can take the average of a bunch of uniform samples in the unit cube to estimate the integral (referred to as *Monte Carlo* integration)
- Let $Z = \frac{\bar{Y} - \mu}{\frac{\sigma}{\sqrt{N}}}$, then by the CLT as $N \to \infty$, Z has standard normal distribution
  - The standard deviation of $\bar{Y} - \mu$ is decreasing at a rate of $\frac{1}{\sqrt{n}}$
  - This means the error is also decreasing at a rate of $\frac{1}{\sqrt{N}} = N^{-\frac{1}{2}}$
- The full steps of Monte Carlo integration:
  1. Make uniform random variables $\begin{bmatrix} x_1^k \\ \vdots \\ x_d^k \end{bmatrix}, k = 1, \cdots, N$ (the sample points)
  2. Evaluate the function at these points $Y^k = f(x_1^k, \cdots, x_d^k)$
  3. The result of the integral is approximately $\frac{1}{N}\sum_{k=1}^{N} Y^k$
- Compare MC's order of $O(N^{-\frac{1}{2}})$ to trapezoidal rule's $O(N^{-\frac{2}{d}})$
  - MC's rate of convergence is unaffected by dimension while trapezoidal rule is highly dependent on dimensionality
  - For lower number of dimensions ($d < 4$) trapezoidal rule is still better
  - For any higher dimensions MC becomes better