Monte Carlo Integration

10.34 Fall 2015

Recall: mean converges with number of samples N

- By Central Limit Theorem of Statistics, as N gets large:

  $\langle y \rangle_N \rightarrow \langle y \rangle_{true}$ with error decreasing as $N^{-1/2}$

  $\sigma^2(\langle y \rangle) \rightarrow \sigma^2(y)/N$

- When we do this average by repeating an experiment, we are sampling from a probability distribution $p_{expt}(y)$. If $p_{expt}(y)$ is Gaussian ("normal distribution") we can generate *synthetic data sets* by random number generator:

```
for i=1:N
        y(i)=sigma*randn + center;
end
```

Two ways to compute <g(y)>

- Suppose we measure y(i) but what we want to know is the expected value of g(y). If we are sampling (experimentally or with synthetic data):

$$\langle g \rangle = (1/N) \sum g(y_i)$$

This is called a "stochastic" method, gives different result each time, but as N gets very large it gets accurate.

- If we know p(y), we could instead compute this integral "deterministically":

$$\langle g \rangle = \int g(y)\, p(y)\, dy$$

probably we would do this using quadrature in uniform intervals, or maybe some adaptive meshing technique (e.g. rewrite it as an ODE and use ode45). These methods would give the same numerical value each time we ran it, no randomness involved (but to get the exact result we still need to go to an infinitely dense mesh).

Any integral can be rewritten this way

- Integrals are averages:

$$\int_a^b f(x)dx = < f > * (b-a)$$

- Can rewrite this as:

$$< f > = \int f(x)p(x)dx$$

$$p(x) = \{H(a) - H(b)\}/(b-a)$$

- So we can compute any integral stochastically by drawing numbers from p(x), which is a uniform distribution from a to b:

```
For i=1:N
  x = a + (b-a)*rand;
  sum_f = sum_f + f(x);
end
 mean_f = sum_f / N;
```

Scaling for Multi-dimensional integrals

- Suppose we have to compute an m-dimensional integral. What way is best?
- For small m, we can use Simpson's rule, 2^m points gives one interval in each dimension. If m=100, need one million function evals. To make it two intervals in each dimension to gain a sig fig, need 4^m points (~10^{12} if m=100). Curse of Dimensionality!
- For large m, better to use stochastic method, error scales as $N^{-1/2}$, can use any number of points. So one million points reduces uncertainty by factor of 1,000. Only need 100x as many samples (~10^8) to gain one more significant figure.
- But for sharply-peaked high-dimensional functions, even 10^8 points may not be enough to have several points in region of the peak.

Metropolis Monte-Carlo

- Often we want to compute integrals of this type (e.g. in statistical mechanics & quantum mechanics, also in turbulent flow simulations)

$$< f >= \iiint f(\underline{x})\, p(\underline{x}) d^m x$$

$$p(\underline{x}) = w(\underline{x}) / \iiint w(\underline{x}) d^m x$$

- Metropolis et al. (1953) invented a method for sampling from p(x) just using w(x)
- When p(x) is sharply peaked and m is large this is much much better than integration techniques based on uniform meshes or uniform distributions.

Metropolis's Method

- See Beers' text pp 353-356!
- Given current x_i need a way to propose a new x_k (e.g. add a random step Δx)
- Metropolis says: accept the step if $w(x_k)>w(x_i)$
- Else

  Compute a random number $0<u<1$ (u = rand)

  Accept the step if $u<w(x_k)/w(x_i)$. Else set $x_k=x_i$.
- Sum_f = Sum_f + $f(x_k)$

What is necessary for Metropolis's method to work?

- The method of proposing new steps must be able to reach all possible x's ("ergodic").
- The overall transition probability =

  (probability of proposing step) * (probability of accepting step)

  must satisfy ***detailed balance***:

$w(x_i)$*(transition probability i → k) =

$w(x_k)$*(transition probability k → i)

10.34 Numerical Methods Applied to Chemical Engineering

Fall 2015