Random numbers & Monte Carlo integration

We have so far considered two methods for "solving" for the partition function and various observables in generalised spin models / in lattice field theories. Both methods have some problems, however:

- the convergence radius of the weak coupling expansion is, strictly speaking, zero. It is an asymptotic expansion,

$$\left| f(x) - \sum_{k=0}^{M} f_k x^k \right| = \mathcal{O}\left(c_M x^{M+1}\right);$$

for fixed M good for sufficiently small x, but for fixed x the approximation gets worse beyond some M. Needs to be used with great care!

- the strong coupling expansion has a finite radius of convergence, for any $\hat{\lambda}$. That radius does not extend beyond the value κ_c, however, which corresponds to a <u>phase transition</u> in spin models, and <u>continuum limit</u> in field theory.

On the other hand,

$$Z = \left\{ \prod_{\vec{x}} \int ds(x) \right\} \exp\left[\sum_{\vec{y}} \left\{ 2\kappa \sum_{i=1}^{d} s(\vec{y}) s(\vec{y}+\hat{\imath}) - s^2(\vec{y}) - \hat{\lambda}\left(s^2(\vec{y})-1\right)^2 \right\} \right]$$

is just a finite-dimensional $(N_1 \cdot N_2 \cdots N_d)$ integral. Why not evaluate it numerically?

Monte Carlo integration

Basic idea: consider a D-dimensional integral,

$$I = \int_0^1 d\phi_0 \int_0^1 d\phi_2 \cdots \int_0^1 d\phi_D \; f(\vec{\phi}) \,,$$

where a simple integration domain has been chosen for simplicity.

Generate N_{meas} random numbers $\vec{\phi}_i$ from a flat distribution. ("measurements").

Then $I \approx \frac{1}{N_{meas}} \sum_{i=1}^{N_{meas}} f(\vec{\phi}_i)$, with an error $\delta I \propto \frac{1}{\sqrt{N_{meas}}} \overset{N_{meas}\to\infty}{\to} 0$.

How does this compare with normal numerical integration methods?

* divide each axis into n evenly spaced intervals.
* total number of points: $N_{meas} = n^D$.
* error: $\propto \frac{1}{n}$ (midpoint rule)
  $\propto \frac{1}{n^2}$ (trapezoidal rule)
  $\propto \frac{1}{n^4}$ (Simpson method)

$\Rightarrow$ Monte Carlo method has a smaller error than Simpson if

$$\frac{1}{\sqrt{N_{meas}}} < \frac{1}{n^4} = \frac{1}{N_{meas}^{\frac{4}{D}}} \,, \quad \text{for } N_{meas} \to \infty$$

$$\Rightarrow D > 8$$

In spin models / lattice field theories,

$$D \geq \sum_{\vec{x}} = N_1 N_2 \cdots N_d \sim 10^6 \quad !$$

Thus MC is the method of choice: numerical integration always outputs some number, and the challenge is to minimise the error with given resources.

Why is the error $\propto \frac{1}{\sqrt{N_{meas}}}$? $\Rightarrow$ central limit theorem

Let us denote $\langle f \rangle \equiv \int d\bar{\phi}\, f(\bar{\phi})$

$\langle f^2 \rangle \equiv \int d\bar{\phi}\, f^2(\bar{\phi})$

$\sigma^2 \equiv \langle f^2 \rangle - \langle f \rangle^2$

We estimate $\langle f \rangle$ by $F_{N_{meas}} \equiv \frac{1}{N_{meas}} \sum_{i=1}^{N_{meas}} f(\bar{\phi}_i)$, where $\bar{\phi}_i$ are random and flat.

What do we get "on the average" (i.e., if we repeated the same many times)?

$$\langle\langle F_{N_{meas}} \rangle\rangle \equiv \int d\bar{\phi}_1 \int d\bar{\phi}_2 \cdots \int d\bar{\phi}_{N_{meas}}\, F_{N_{meas}} = \langle f \rangle .$$

What is the error made?

$$\sigma^2_{N_{meas}} = \langle\langle F^2_{N_{meas}} \rangle\rangle - \langle\langle F_{N_{meas}} \rangle\rangle^2$$

$$= \int d\bar{\phi}_1 \int d\bar{\phi}_2 \cdots \int d\bar{\phi}_{N_{meas}}\, \frac{1}{N^2_{meas}} \sum_{i,j=1}^{N_{meas}} f(\bar{\phi}_i) f(\bar{\phi}_j) - \langle f \rangle^2$$

$$= \frac{1}{N_{meas}} \langle f^2 \rangle + \frac{1}{N^2_{meas}} N_{meas}(N_{meas}-1) \langle f \rangle^2 - \langle f \rangle^2$$

$$= \frac{1}{N_{meas}} \sigma^2$$

This is called the "1-σ error": $\sigma_{N_{meas}} = \frac{\sigma}{\sqrt{N_{meas}}}$.

Refinements:

(a) In the real world, we do not know $\langle f \rangle$, $\langle f^2 \rangle$, but they are estimated by $\frac{1}{N_{meas}} \sum_i f(\bar{\phi}_i)$, $\frac{1}{N_{meas}} \sum_i f^2(\bar{\phi}_i)$.

Then the correct unbiased error estimate is

$$\sigma_{N_{meas}} = \sqrt{\frac{\frac{1}{N_{meas}} \sum_i f^2(\bar{\phi}_i) - \left(\frac{1}{N_{meas}} \sum_i f(\bar{\phi}_i) \right)^2}{N_{meas} - 1}} .$$

(b) We will return to better methods (jackknife, bootstrap) later on.

Random numbers

Monte Carlo integration relies on random numbers. The procedure has to be reproducible, however ⇒ use pseudorandom numbers, which are actually deterministic.

[genuine random numbers are needed, e.g., in cryptography]

The generic procedure:
(1) give seed.
(2) generate integers from it.
(3) convert to a floating point number $\in [0,1)$ by dividing by a constant.

Properties required:
(1) repeatable (= deterministic)
(2) uniform
(3) non-correlated (never quite!)
⇒ long period, longer than the simulation
(4) independent of seed (test!)
(5) fast
(6) portable (= works on any computer)

An example: LCG = linear congruential generator (Lehmer 1948)

$$X_{i+1} = (aX_i + c) \bmod m \quad , \quad a, c, m \in \mathbb{N}$$

And divide then by m.
Period at most m. The constants a, c have to be chosen carefully!

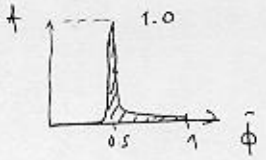
Eg. UNIX drand48 ( ): $a = 5\text{DEECE66D}_{16}$, $c = B_{16}$, $m = 2^{48}$
NAG $a = 13^{13}$, $c = 0$, $m = 2^{59}$

Other types: "Lagged Fibonacci"; "Mersenne twister"; RANMAR; RANLUX; CMRG
Of these RANLUX (arxiv.org/abs/hep-lat/9309020) has mathematically proven properties.

References:
- Press, Teukolsky, Vetterling, Flannery, Numerical Recipes [www.nr.com]
- NAG libraries [www.nag.co.uk]
- D. Knuth, The Art of Computer Programming, Vol. 2
- www.iro.umontreal.ca/~simardr/random.html
- random.mat.sbg.ac.at
- www.physics.helsinki.fi/~vatthulai/rngs.html
- www.physics.helsinki.fi/~rummukai/simu/...

The basic idea of importance sampling

Suppose f is sharply "localised":



Then an even distribution of points for the $\bar{\phi}$-integral is not efficient, and leads to a big error.

Example: (a) $N_{meas} = 10$, $\bar{\phi}_i = 0.5$, $f(0.5) = 1$, other f's $= 0$ $\Rightarrow$ $\delta_{N_{meas}} = \sqrt{\frac{\frac{1}{10}\cdot 1 - \left(\frac{1}{10}\right)^2}{9}} = 0.1$

(b) $N_{meas} = 10$, all $\bar{\phi}_i$'s around 0.5 $\Rightarrow$ $\delta_{N_{meas}} = 0.0$

Integral itself $= 0.1 \times 1.0 = 0.1$! $\Rightarrow$ method (a) is useless!

Importance sampling allows to choose points close to the important region.

$$I = \int d\bar{\phi}\, f(\bar{\phi}) = \int d\bar{\phi}\, p(\bar{\phi})\, \frac{f(\bar{\phi})}{p(\bar{\phi})}$$

Suppose we find a variable $\breve{\phi}$ such that

$$d\breve{\phi} = p(\bar{\phi})\, d\bar{\phi}$$

Then $$I = \int d\breve{\phi}\, \frac{f(\bar{\phi}(\breve{\phi}))}{p(\bar{\phi}(\breve{\phi}))} = \frac{1}{N_{meas}} \sum_{\breve{\phi}_i} \frac{f(\bar{\phi}(\breve{\phi}))}{p(\bar{\phi}(\breve{\phi}))}$$

If $p(\bar{\phi}) \approx f(\bar{\phi})$, the new distribution is flatter, and error is smaller.

How to implement this in practice?

* generate configurations $\bar{\phi}$ from a flat distribution.
* "accept" them with probability $p(\bar{\phi})$, "reject" with $1 - p(\bar{\phi})$.
* carry out "measurements" with accepted $\bar{\phi}$ only.
* once one finds a $\bar{\phi}$ in an important region, should only change it by a small amount!

A proper implementation leads to "update algorithms".