Computational Physics (6810): Session 13

Dick Furnstahl
Nuclear Theory Group
OSU Physics Department

April 14, 2017

6810 Endgame

Various recaps and followups

Random stuff (like RNGs :)

Session 13 stuff
6810 Computational Physics Endgame

- April 21 is our last class period!
- Projects are due at the end of Monday, May 1
  - Create a Project subfolder of your BuckeyeBox folder
  - Include codes, makefiles, plots, etc. (like for problem sets)
  - Include an explanation of how it all fits together in a separate file or in comments of the codes.
  - It is recommended that you turn in something earlier than the due date to get feedback while there’s time for iteration.
- Session guides and homework will be accepted until 5pm on Thursday, April 27
  - Remember that you can always upgrade a check
  - Get in some version (even if incomplete) soon!
Which one is not really random?
Gaussian or Normal distributions

- Gaussian is **fully** characterized by mean $\mu$ and variance $\sigma^2$

$$p(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Standard deviation $\sigma$ is square root of variance
- About 2/3 of numbers in $[-\sigma, +\sigma]$ “within 1 sigma”
- PDG requires a “5$\sigma$ signal” to declare a new particle is found
Comments on probability distribution functions (PDFs)

- Two common ones are uniform and gaussian [think histograms!]
  - Always normalized: $\int_{-\infty}^{\infty} p(x) \, dx = 1$ (total probability is one)
  - Uniform: $p(x) = \frac{1}{b-a} \theta(x - a) \theta(b - x)$
Comments on probability distribution functions (PDFs)

- Two common ones are uniform and gaussian [think histograms!]
  - Always normalized: \( \int_{-\infty}^{\infty} p(x) \, dx = 1 \) (total probability is one)
  - Uniform: \( p(x) = \frac{1}{b-a} \theta(x - a) \theta(b - x) \)

- Gaussian is \textit{fully} characterized by mean \( \mu \) and variance \( \sigma^2 \)
  \[
p(x) = \frac{1}{2\pi\sigma^2} e^{-(x-\mu)^2/2\sigma^2}
  \]
  - Standard deviation \( \sigma \) is square root of variance
  - About 2/3 of numbers in \([-\sigma, +\sigma]\) “within 1 sigma”
  - PDG requires a “5\(\sigma\) signal” to declare a new particle is found
Comments on probability distribution functions (PDFs)

- Two common ones are uniform and gaussian [think histograms!]
  - Always normalized: \( \int_{-\infty}^{\infty} p(x) \, dx = 1 \) (total probability is one)
  - Uniform: \( p(x) = \frac{1}{b-a} \theta(x-a) \theta(b-x) \)

- Gaussian is fully characterized by mean \( \mu \) and variance \( \sigma^2 \)
  \[
p(x) = \frac{1}{2\pi\sigma^2} e^{-(x-\mu)^2/2\sigma^2}
  \]

- Standard deviation \( \sigma \) is square root of variance
- About 2/3 of numbers in \([-\sigma, +\sigma]\) “within 1 sigma”
- PDG requires a “5\(\sigma\) signal” to declare a new particle is found

- 2D random walk and \( \sqrt{N} \): How far away after \( N \) steps? (label by \( i \))
  \[
  R_{\text{avg}}^2 \approx \langle (\Delta x_1 + \Delta x_2 + \cdots + \Delta x_N)^2 + (\Delta y_1 + \Delta y_2 + \cdots + \Delta y_N)^2 \rangle
  \]
  - This is not one walk, but the average of many (note \( \langle \rangle \)’s)
  - \( \Delta x_i \) and \( \Delta y_i \) chosen by \( p(x) \) with means \( \langle \Delta x_i \rangle = \langle \Delta y_i \rangle = 0 \)
  - Uncorrelated steps means \( \langle \Delta x_i \Delta x_j \rangle = \langle \Delta y_i \Delta y_j \rangle = 0 \) if \( i \neq j \)
  \[
  \implies R_{\text{avg}}^2 \approx N \langle \Delta x_i^2 + \Delta y_i^2 \rangle \equiv N \langle r^2 \rangle \implies R_{\text{rms}} \approx \sqrt{Nr_{\text{rms}}}
  \]

- Applies generally to processes with combined random errors
Scaling of fluctuations with $N$

Uniform Random Number Distributions

Four times as many points on top. Scaling of fluctuations?
Scaling of fluctuations with $N$

Four times as many points on top. Scaling of fluctuations?
answer: proportional to $\sqrt{N}$, so twice as large on top

What is the *relative* scaling of the fluctuations with $N$?
Scaling of fluctuations with $N$

Four times as many points on top. Scaling of fluctuations?
answer: proportional to $\sqrt{N}$, so twice as large on top

What is the relative scaling of the fluctuations with $N$?
answer: $\frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}$
Things to watch for in Session 12 and beyond

- For Monte Carlo methods, we need a set of random numbers
  - *uncorrelated* means we can’t predict $x_{i+1}$ given $x_1, x_2, \cdots, x_i$
- Doesn’t have to be uniform
  - the histogram of random numbers will approach the shape of the corresponding probability distribution function (PDF)
- We’ll use *pseudo*-random numbers from GSL random number generators (rng)
  - trade-offs between period and speed
  - see Session 12 notes for tests (e.g., using your eye) and Session 13 notes for pitfalls!
  - You need to *seed* the rng, or you’ll get the same numbers!
Things to watch for in Session 12 and beyond (cont.)

- Random walks in two dimensions: $N$ steps
  - Be able to derive standard deviation distance $R \approx \sqrt{Nr_{\text{rms}}}$
  - Remember how to find the average of a function in an interval:

  \[
  \langle f \rangle = \frac{1}{b-a} \int_a^b f(x) \, dx \quad \Rightarrow \quad \langle r^2 \rangle = \frac{1}{(b-a)^2} \int_a^b \int_a^b dy \, (x^2+y^2)
  \]

- (Crude) approximation of an integral $I$

  \[
  \langle f \rangle = \frac{1}{b-a} \int_a^b f(x) \, dx \approx \frac{1}{N} \sum_{i=1}^N f(x_i) \quad \Rightarrow \quad I = \int_a^b f(x) \, dx \approx \frac{b-a}{N} \sum_{i=1}^N f(x_i)
  \]

- Better approximation: sample according to an appropriate pdf

  \[
  I = \int_a^b dx \, f(x) = \int_a^b dx \, w(x) \frac{f(x)}{w(x)} = \left\langle \frac{f}{w} \right\rangle_w \approx \frac{1}{N} \sum_{i=1}^N \frac{f(x_i)}{w(x_i)}
  \]

  \[\Rightarrow\] Sample the integrand (i.e., choose points) where big or steep
Averaging functions over probability distributions

\[ \langle f(x) \rangle \equiv \int f(x) P(x) \, dx \quad \text{where} \quad \int P(x) \, dx = 1 \quad \text{(usually from} \ -\infty \ \text{to} \ +\infty) \]

generalizes to

\[ \langle g(x, y) \rangle \equiv \int \int g(x, y) P_2(x, y) \, dx \, dy \quad \text{where} \quad \int \int P_2(x, y) \, dx \, dy = 1 \]
Averaging functions over probability distributions

\[ \langle f(x) \rangle \equiv \int f(x) P(x) \, dx \] where \[ \int P(x) \, dx = 1 \] (usually from \(-\infty\) to \(+\infty\))

generalizes to

\[ \langle g(x, y) \rangle \equiv \int \int g(x, y) P_2(x, y) \, dx \, dy \] where \[ \int \int P_2(x, y) \, dx \, dy = 1 \]

\[ \langle r^2 \rangle = \langle x^2 + y^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle = 2\langle x^2 \rangle \]

Here we have special cases of \( P_2 \) and \( g \):

\[ P_2(x, y) = P(x)P(y) \quad \text{and} \quad g(x, y) = f(x) + f(y) \]
Averaging functions over probability distributions

\[ \langle f(x) \rangle \equiv \int f(x) P(x) \, dx \quad \text{where} \quad \int P(x) \, dx = 1 \quad (\text{usually from } -\infty \text{ to } +\infty) \]

generalizes to

\[ \langle g(x, y) \rangle \equiv \int\int g(x, y) P_2(x, y) \, dx \, dy \quad \text{where} \quad \int\int P_2(x, y) \, dx \, dy = 1 \]

\[ \langle r^2 \rangle = \langle x^2 + y^2 \rangle = \langle x^2 \rangle + \langle y^2 \rangle = 2\langle x^2 \rangle \]

Here we have special cases of \( P_2 \) and \( g \):

\[ P_2(x, y) = P(x)P(y) \quad \text{and} \quad g(x, y) = f(x) + f(y) \]

\[ \langle f(x) + f(y) \rangle = \int\int [f(x) + f(y)] P(x)P(y) \, dx \, dy \]

\[ = \left[ \int f(x) P(x) \, dx \right] \times \left[ \int P(y) \, dy \right] + \left[ \int P(x) \, dx \right] \times \left[ \int f(y) P(y) \, dy \right] \]

\[ = \int f(x) P(x) \, dx + \int f(y) P(y) \, dy = \langle f(x) \rangle + \langle f(y) \rangle \checkmark \]
Ways to do Importance Sampling

- In general, identify a pdf $P(x)$ in your integrand:

$$\int_a^b f(x) \, dx \longrightarrow \int_a^b g(x) P(x) \, dx \approx \frac{1}{N} \sum_{i=1}^{N} g(\tilde{x}^{(i)})$$

- where the points $\{\tilde{x}^{(i)}\}$, $i = 1, N$, are sampled from $P(x)$

- Note that $P(x)$ does not appear explicitly in the sum

This is called “Importance Sampling”

- Pick a distribution $P(x)$ that matches the integrand as best as possible (e.g., make $g(x)$ as close to constant as possible)

Ways to determine the $\tilde{x}$’s:

1. Choose an explicit $P(x)$ if you know how to sample it (e.g., a Gaussian distribution)

2. Determine $P(x)$ adaptively. E.g., `vegas` in GSL

3. Build up a set of configurations $\{\tilde{x}^{(i)}\}$ via a Markov chain (e.g., via the Metropolis algorithm or some variation)
First pass at statistical mechanics basics

- The state of our (approximated) system is specified by a *configuration*
  - For example, for an \( N \)-particle quantum system, it could be where each particle is located (plus other quantum #’s)
  - For the Ising model, it is whether each “spin” on the lattice (which could be in any number of dimensions) is up or down
  - Our approximation to the system is such that specifying a configuration takes a finite amount of data but the number of *possible* configurations can still be (and usually is!) enormous.

- The *partition function* tells us what we need for thermodynamics
  - E.g., the expectation value of the energy or of an observable like the pressure or magnetization
  - We can express it two equal ways (canonical ensemble here)

\[
Z = \sum_i \frac{e^{-E^{(i)}/k_B T}}{\Omega(E^{(i)})} = \sum_j \frac{e^{-E_j/k_B T}}{\Omega(E_j)}
\]

- The first sum is over all configurations, labeled by \( i \)
- The second sum is over all energies, labeled by \( j \)
Example of one-dimensional Ising model with $N$ spins

- Specify a configuration by whether each of $N$ spins is up or down:

\[
\begin{align*}
\uparrow & \quad \downarrow & \quad \uparrow & \quad \uparrow & \quad \downarrow & \quad \cdots & \quad \uparrow \\
0 & \quad 1 & \quad 2 & \quad 3 & \quad 4 & \quad 5 & \quad N-1
\end{align*}
\]

- Suppose the Hamiltonian specified the energy as: $-J \sum_{\langle ij \rangle} S_i S_j$

  - Here $J$ is a constant with dimensions of energy
  - $\langle ij \rangle$ means all adjacent spins ("nearest neighbors")
  - $S_i = 1$ for $\uparrow$ and $S_i = -1$ for $\downarrow$ (e.g., so $S_i S_j = 1$ if both down)

- Consider $N = 3$ with periodic boundary conditions: states and energies

<table>
<thead>
<tr>
<th>config. $i$</th>
<th>state</th>
<th>energy (with $J = 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>↓↓↓</td>
<td>$E^{(0)} = -(+1 + 1 + 1) = -3$</td>
</tr>
<tr>
<td>1</td>
<td>↓↓↑</td>
<td>$E^{(1)} = -(+1 - 1 - 1) = +1$</td>
</tr>
<tr>
<td>2</td>
<td>↓↑↓</td>
<td>$E^{(2)} = -(+1 + 1 + 1) = +1$</td>
</tr>
<tr>
<td>3</td>
<td>↓↑↑</td>
<td>$E^{(3)} = -(+1 + 1 + 1) = +1$</td>
</tr>
<tr>
<td>4</td>
<td>↑↓↓</td>
<td>$E^{(4)} = -(+1 + 1 + 1) = +1$</td>
</tr>
<tr>
<td>5</td>
<td>↑↓↑</td>
<td>$E^{(5)} = -(+1 + 1 + 1) = +1$</td>
</tr>
<tr>
<td>6</td>
<td>↑↑↓</td>
<td>$E^{(6)} = -(+1 + 1 + 1) = +1$</td>
</tr>
<tr>
<td>7</td>
<td>↑↑↑</td>
<td>$E^{(7)} = -(+1 + 1 + 1) = -3$</td>
</tr>
</tbody>
</table>

There are 7 different configurations but $E_0 = -3$ and $E_1 = +1$ are the only two choices for $E_j$.

Find $Z$ by summing over configurations $i$ or energies $E_j$.
More statistical physics in a nutshell (I)

Maximizing the Free Energy (not the energy!)

Return to partition function and probability $P(E)$ of finding energy $E$:

$$Z = \sum_{\text{all } i} e^{-E^{(i)}/k_B T} = \sum_{E_j} (\# \text{ of } E_j \text{s}) e^{-E_j/k_B T} \equiv \sum_{E_j} \Omega(E_j) e^{-E_j/k_B T}$$

Here $P(E) \propto \Omega(E) e^{-E/k_B T}$. The system will maximize $E$. How?

- Move $\Omega(E)$ to exponent, as $e^{-\alpha}$ maximized when $\alpha$ minimized
- $\Omega = e^{\ln \Omega} \implies P(E) \propto e^{-(E-TS)/k_B T}$ with $S = k_B \ln \Omega(E)$
- What is the minimized quantity at fixed $T$? $E - TS$
- Note the tradeoff between $E$ and $TS$!
More statistical physics in a nutshell (II)

Temperature and Entropy

- Consider two subsystems with energies $E$ and $E'$ and fixed $E_{\text{tot}} = E + E'$
- Exchange energy back and forth, reach equilibrium. Meaning what? Static?
- Define $\Omega(E)$ and $\Omega(E')$ to be the number of configurations with $E$ and $E'$ (i.e., “microstates”). How many total ways? $\Omega(E) \times \Omega(E') = \Omega(E)\Omega(E_{\text{tot}} - E)$
- Basic principle: any accessible state (satisfies $E + E' = E_{\text{tot}}$) is equally likely. $\implies$ probability of particular $E$ is total number with that $E$ divided by grand total
- So maximize $\Omega(E)\Omega(E')$ to maximize probability $\implies$ Why is this equilibrium?
- Because $\Omega(E) = e^{S(E)/k_B}$, maximize $S(E) + S(E')$. 2nd Law of thermo?
- Set derivative to zero and use chain rule with $dE'/dE = -1$ (why?):

$$\frac{d}{dE}(S(E) + S(E')) = 0 = \frac{dS(E)}{dE} + \frac{dS(E')}{dE'} \frac{dE'}{dE}$$

- So equilibrium means $\frac{dS(E)}{dE} = \frac{dS(E')}{dE'}$
- We use this to define the temperature $T = \frac{dS(E)}{dE}$!
More statistical physics in a nutshell (III)

Probability and Boltzmann factors

- Find the probability of energy $P(E)$ for one of two subsystems that is very much smaller than the other, so $E \ll E_{\text{tot}}$

- Ratio of probabilities is ratio of total number of possible states

\[
\frac{P(E_1)}{P(E_2)} = \frac{\Omega(E_1)\Omega(E_{\text{tot}} - E_1)}{\Omega(E_2)\Omega(E_{\text{tot}} - E_2)} = \frac{\Omega(E_1)e^{\frac{1}{k_B}S(E_{\text{tot}} - E_1)}}{\Omega(E_2)e^{\frac{1}{k_B}S(E_{\text{tot}} - E_2)}}
\]

- Now Taylor expand (of course!): $S(E_{\text{tot}} - E_1) \approx S(E_{\text{tot}}) - E_1 \left. \frac{dS}{dE} \right|_{E_{\text{tot}}}$

- But $\frac{dS}{dE} = \frac{1}{T}$ and $e^{S(E_{\text{tot}})/k_B}$ factors cancel, leaving

\[
\frac{P(E_1)}{P(E_2)} \approx \frac{\Omega(E_1)e^{-E_1/k_B T}}{\Omega(E_2)e^{-E_2/k_B T}}
\]

- So $P(E) \propto \Omega(E)e^{-E/k_B T}$. How to normalize $P(E)$?

\[
\sum_i \Omega(E_i)e^{-E_i/k_B T} = 1 \implies P(E) = \frac{1}{Z}\Omega(E)e^{-E/k_B T}
\]
Example of one-dimensional Ising model with 20 spins

Energy Distributions

random sampling
exact T=10.
Metropolis T=10.
exact T=1.0
Metropolis T=1.0

Sat Mar 26 08:48:39 2016
Example of one-dimensional Ising model with 20 spins

![Energy Distributions](image.png)

random sampling
exact T=10.
Metropolis T=10.
exact T=1.0
Metropolis T=1.0
Metropolis small T=1.0

Wed Apr 20 11:37:10 2016
random sampling
exact T=10.
Metropolis T=10.
exact T=1.0
Metropolis T=1.0
Metropolis small T=1.0

Dick Furnstahl  6810: Session 13
Example of one-dimensional Ising model with 20 spins

Energy Distributions

random sampling
exact T=10.
Metropolis T=10.
exact T=1.0
Metropolis T=1.0
Metropolis small T=1.0


Dick Furnstahl
Two-D Ising model: Equilibration and cooling

2D Ferromagnetic Ising Model Energy vs. Time using Monte Carlo

energy E vs. time t

Sat Mar 26 08:53:38 2016

Dick Furnstahl