Project 1, Fys 4411

Karl R. Leikanger

February 24, 2012

Abstract

1 Quantum dots

Our system of choice

2 The Monte Carlo method

2.1 Introduction

Monte Carlo (MC) methods includes a wide range of computational methods for solving integrals. The main idea is to sample the integrand $f(\vec{r})$ in random points $\{\vec{r}_i\}$, and finding an approximate solution to the integral by doing statistical analysis on the sampling data. The statistical error will give us an immediate idea of the precicion of the result.

In general, any integral can be written on the form

$$\int_{\mathcal{D}} f(\vec{r})d\vec{r} = \int_{\mathcal{D}} P(\vec{r})g(\vec{r})d\vec{r}, \tag{1}$$

where $P(\vec{r})$ is a probability distribution. The integral can now be approximated by the sum

$$\int_{\mathcal{D}} f(\vec{r})d\vec{r} \approx \frac{1}{N}\sum_{i=1}^{N} g(\vec{r}_i), \tag{2}$$

where $\{\vec{r}_i\}$ is a set of random numbers drawn from the statistical distribution $P(\vec{r})$.

The simplest example, possible whenever the region of the of the integral is finite, is when P is a uniform distribution bounded by the integration limits. Then $P(\vec{r}) = 1$ and $g(\vec{r}) = f(\mathbf{r})$. However, for many integrals this method will converge slowly and will require a very large number of samples to give accurate results. By choosing a smart $P(\vec{r})$ that is large in the regions of interest, more samples will be drawn from the important regions, often leading to a more rapid convergence.

2.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) provides us with a method for drawing the correct samples $\{\vec{r}_i\}$ from P. Instead of generating random samples we use a markow process to model P. The Markov chain is caracterized by a transition kernel $q(\vec{x}|\vec{y})$ which tells us the probability of a transition from $\vec{x}$ to $\vec{y}$. Note that $q(\vec{x}|\vec{y})$ is a probabitity density distribution of the variable $\vec{y}$. We will show that we can sample P by applying any transition kernel $q(\vec{x}, \vec{y})$ to which P is invariant.

$$\int_{\mathcal{D}} d\vec{r}_A\, P(\vec{r}_A)q(\vec{r}_A|\vec{r}_B) = P(\vec{r}_B) \tag{3}$$

Our Markow chain is analouge to the description of a set of random walkers, where the density of walkers P_w is stationary and proportional to P.

If we sample each point on the track of a walker, this set will converge to $P_w \propto P$. In the limit of an infinite number of walkers, the population of walkers in any point $\vec{r}_A$ must be constant,

$$\sum_n P_w(\vec{r}_A)q(\vec{r}_A|\vec{r}_n) = \sum_n P_w(\vec{r}_n)q(\vec{r}_n|\vec{r}_A) \tag{4}$$

This equation is equivalent to (3) since they both states that the density is stationary. One are free to choose any equation for the trajectories of the random walkers as long as this equation is fulfilled. One obvious solution is to assume detailed balance,

$$P_w(\vec{r}_A)q(\vec{r}_A|\vec{r}_B) = P_w(\vec{r}_B)q(\vec{r}_B|\vec{r}_A). \tag{5}$$

This equation is the basis for finding the transition kernels that is used in the metropolis-Hastings algorithm.

2.3 Metropolis-Hasting algorithm

We will now outline a general algorithm for the movement of the random walkers that fullfills (5).

We introduce new scalar field to our equation $\gamma(\vec{r}_A, \vec{r}_B)$, which is a measure of the probability of accepting a move from $\vec{r}_A$ to $\vec{r}_B$. We split our transition kernel in two parts

$$q(\vec{r}_A|\vec{r}_b) \rightarrow \gamma(\vec{r}_A|\vec{r}_b).g(\vec{r}_A|\vec{r}_b) \tag{6}$$

We set $\gamma(\vec{r}_B|\vec{r}_A) = 1$ when

$$P_w(\vec{r}_A)g(\vec{r}_A|\vec{r}_B) > P_w(\vec{r}_B)g(\vec{r}_B|\vec{r}_A). \tag{7}$$

Then, by inserting this into (5), we get

$$\gamma(\vec{r}_A|\vec{r}_B) = min\left(\frac{P_w(\vec{r}_B)g(\vec{r}_B|\vec{r}_A)}{P_w(\vec{r}_A)g(\vec{r}_A|\vec{r}_B)}, 1\right) \tag{8}$$

By making this choice for γ, (5) will be fulfilled and the sequence of $\{r_i\}$ will sample P. This rule preserves the relative provability of the transitions between r_A and r_B, and therefore preserves the density of walkers in each point. Note that only the relative ratio between the left and the right hand side of (eq) are of interest to us. Therefore, P does not have to be normalized.

These equations are implemented in the well known Metropolis-Hastings algorithm as follows:

loop
  $\vec{r}_B \leftarrow$ random from distribution $g(\vec{r}_T|\vec{r}_B)$.
  $\mathcal{X} \leftarrow$ random uniform. ($\mathcal{X} \in [0, 1]$)
  Calculate $\gamma(\vec{r}_n|\vec{r}_T)$).
  if $\mathcal{X} \leq \gamma$ **then**
    $\vec{r}_{n+1} \leftarrow \vec{r}_T$
  else
    $\vec{r}_{n+1} \leftarrow \vec{r}_n$
  end if
  $n \leftarrow n + 1$
end loop

Brute force metropolis sampling: transition kernel $g(x|y)$ any distribution symmetric around x.
Metropolis-Hastings: $g(x|y)$, trajectories gouverned by Langevin dynamics. (trajectories of walkers: the solution to ... equation).

For this algorithm to work, it is essential that the system of random walkers is ergodetic. That means that all points must be accessible for every single walker. Otherwise, some parts of the integration domain would not be sampled.

An more sophisticated way of sampling P is the Metropolis - Hastings algorithm.

We will outlining the Metropolis algorithm.
Marcow chains and random walkers.
Metropolis hastings.

For many integrals A possible problem doing importance sampling on the .. is to choose an .. that P has to be normalized. In quantum monte carlo the metropolis algorithm...

2.4 Quantum monte Carlo

In the Heisenberg image, the expectation values for any observable $\hat{O}$ can be expressed on the general form

$$\langle O(t)\rangle = \int d\vec{r}\Psi(\vec{r})\hat{O}(t)\Psi(\vec{r})^*, \tag{9}$$

where $\Psi(\vec{r}) = \sum_i \psi_i(\vec{r})$, for some basis $\{\vec{\psi_i}\}$. The integral can also be written

$$\langle O(t)\rangle = \int |\Psi(\vec{r})|^2 \frac{\hat{O}(t)\Psi(\vec{r})^\dagger}{\Psi(\vec{r})^\dagger}, \tag{10}$$

Since $|\Psi(\vec{r})|^2$ is a probability density distribution we see that

$$\langle O(t)\rangle \approx \sum_{i=1}^{N} \frac{\hat{O}(t)\Psi(\vec{r}_i)^\dagger}{\Psi(\vec{r}_i)^\dagger}, \tag{11}$$

where the set of numbers $\{\vec{r}_i\}$ are drawn from $|\Psi|^2$. This sum can be sampled using the metropolis algorithm with the metropolis ratio $|\Psi(r_i)|^2/|\Psi(r_{i+1})|^2$. Note that Ψ do not need to ne normalized in the above equations. This saves us a great deal of work since the normalization of most wavefunctions would have to be done numerically for every new trial function.

Sohw that the variance is exactly 0 for the correct wf.

For our system, $\hat{O} \leftarrow \hat{H}$, and $\{\psi_i\}$ is the lowest energy eigenstates of the system. Since $[\hat{H}, \int dr\psi_i(\vec{r})\psi_i(\vec{r})] = 0$, we can assume that the the basis functions $\{\psi_i\}$ are real valued without losing generality.

The quantum force: We will assume that

2.5 The statistical analysis of the data

Contents