Maximum entropy method for determination of density of states

Enzhi Li

November 26, 2016

1 Maximum entropy method

Maximum entropy method is widely used to extract the density of states (DOS) from the knowledge of Matsubara Green function, whether be it in imaginary time space or imaginary frequency space. It is a method that is based upon Bayesian analysis, and can give us the best possible inference given a finite amount of information. A detailed review of this method is already available in [1]. Here, I will give only a brief introduction

In this note, I will determine the DOS from the imaginary frequency Matsubara Green function. The spectral function, which is proportional to the DOS, is related with the Green function through

$$\begin{aligned} G(i\omega_n) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{A(\omega)}{i\omega_n - \omega} \\ &:= KA(i\omega_n) \end{aligned} \tag{1}$$

Here, $\omega_n = \frac{(2n+1)\pi}{\beta}$ is a discrete variable, and ω is a continuous variable, K is the kernel function, and KA is a short-hand notation for the integration over ω. According to [1], we define a quantity Q,

$$Q = \alpha S - \frac{1}{2}\chi^2. \tag{2}$$

Here, α is a real number that is introduced for annealing purpose, S is the entropy, and χ^2 is the difference between the actual value of Green function and the expected value of Green

function. The entropy is defined as

$$S = -\int_{-\infty}^{\infty} d\omega A(\omega) \log \frac{A(\omega)}{D(\omega)}, \tag{3}$$

where, $D(\omega)$ is the default model, and χ^2 is defined as

$$\begin{aligned}
\chi^2 &= \sum_{n,m} \Big(G_R(i\omega_n) - \int_{-\infty}^{\infty} d\omega K_R(\omega_n,\omega)A(\omega)\Big)(C_R^{-1})_{nm} \\
&\times \Big(G_R(i\omega_m) - \int_{-\infty}^{\infty} d\omega K_R(\omega_m,\omega)A(\omega)\Big) \\
&+ \sum_{n,m} \Big(G_I(i\omega_n) - \int_{-\infty}^{\infty} d\omega K_I(\omega_n,\omega)A(\omega)\Big)(C_I^{-1})_{nm} \\
&\times \Big(G_I(i\omega_m) - \int_{-\infty}^{\infty} d\omega K_I(\omega_m,\omega)A(\omega)\Big) \\
&:= \sum_{n,m} \Big(G_R - K_R A\Big)_n (C_R^{-1})_{nm} \Big(G_R - K_R A\Big)_m \\
&+ \sum_{n,m} \Big(G_I - K_I A\Big)_n (C_I^{-1})_{nm} \Big(G_I - K_I A\Big)_m
\end{aligned} \tag{4}$$

C_R, C_I are the covariance matrix of the real and imaginary part of the Green function. Here, we have defined the quantities

$$\begin{aligned}
K_R(\omega_n,\omega) &= -\frac{1}{2\pi}\frac{\omega}{\omega^2+\omega_n^2} \\
K_I(\omega_n,\omega) &= -\frac{1}{2\pi}\frac{\omega_n}{\omega^2+\omega_n^2},
\end{aligned} \tag{5}$$

and

$$\begin{aligned}
\Big(K_R A\Big)_n &= \int_{-\infty}^{\infty} d\omega K_R(\omega_n,\omega)A(\omega) \\
\Big(K_I A\Big)_n &= \int_{-\infty}^{\infty} d\omega K_I(\omega_n,\omega)A(\omega)
\end{aligned} \tag{6}$$

In the maximum entropy method, we will calculate the extremal values for Q, and the extremum is obtained by setting the gradient of Q to zero, that is, $\nabla_A Q = 0$. The gradient

of Q with respect to $A(\omega)$ is

$$\begin{aligned} f(A(\omega)) &= \frac{\delta Q}{\delta A(\omega)} \\ &= -\alpha\Big(1+\log\frac{A(\omega)}{D(\omega)}\Big) \\ &+ \sum_{nm} K_R(\omega_n,\omega)(C_R^{-1})_{nm}\Big(G_R - K_R A\Big)_m \\ &+ \sum_{nm} K_I(\omega_n,\omega)(C_I^{-1})_{nm}\Big(G_I - K_I A\Big)_m \end{aligned} \tag{7}$$

We will use Newton's method to calculate the roots of the equation $f(A(\omega))=0$. In Newton's method, we are to solve the iteration equation

$$A_{n+1}(\omega) = A_n(\omega) - \sum_{\omega'} J^{-1}_{\omega,\omega'} f(\omega'). \tag{8}$$

Here, we have defined the Jacobian matrix as

$$\begin{aligned} J_{\omega,\omega'} &= \frac{\delta f(A(\omega))}{\delta A(\omega')} \\ &= -\alpha\delta(\omega-\omega')\frac{1}{A(\omega)} \\ &- \sum_{nm} K_R(\omega_n,\omega)(C_R^{-1})_{nm}K_R(\omega_m,\omega') \\ &- \sum_{nm} K_I(\omega_n,\omega)(C_I^{-1})_{nm}K_I(\omega_m,\omega') \end{aligned} \tag{9}$$

When I run the program, I start from a very large α value. It can be seen from the definition of Q that when $\alpha\to\infty$, the extremal values of Q will coincide with the extremal values of S. Setting the gradient of entropy to zero, we have

$$\frac{\delta S}{\delta A(\omega)} = -1 - A(\omega)\log\frac{A(\omega)}{D(\omega)} = 0 \tag{10}$$

Solving the above equation, we have $A(\omega)=e^{-1}D(\omega)$. Therefore, when α is very large, the spectral function we have should be very similar to the default model. I have used the Gaussian as the default model, and run the program by initializing the initial spectral function at small α from the calculated spectral function at large α. Running the program from large α to small α is called the annealing process.

2 DOS results

2.1 Test case

I am going to use the python program that I developed to calculate the DOS from the knowledge of imaginary frequency Green function. In order to test if the program is correct, I assume we have a prior a spectral function, which is

$$A(\omega) = \frac{1}{2}N(\omega, -1.2, 0.5) + \frac{1}{2}N(\omega, 1.2, 0.4). \tag{11}$$

Here, $N(\omega, \mu, \sigma)$ means a normal distribution with mean value μ and standard deviation σ. From this spectral function, I can calculate the imaginary frequency Green function. I further added Gaussian noise to the Green function, and from the Green function, I extracted the spectral function, and compare the calculated spectral function with the original spectral function. The comparison is shown in Fig. 2.1.

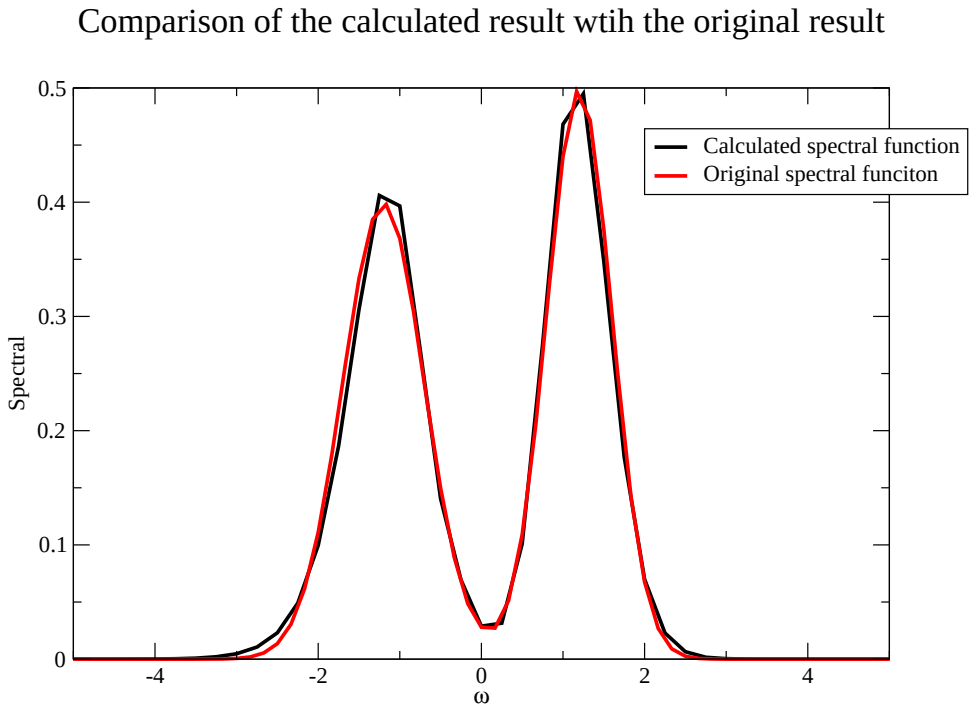

Figure 2.1: Comparison of the calculated spectral function with the original known spectral function.

It can be seen that the calculated result is almost the same as the original spectral function. Therefore, the program can generate the correct spectral function from the knowledge of the imaginary frequency Green function.

References

[1] Mark Jarrell and James E Gubernatis. Bayesian inference and the analytic continuation of imaginary-time quantum monte carlo data. *Physics Reports*, 269(3):133–195, 1996.