Variational Method

Recall that in quantum mechanics, the ground state can be defined variationally as

$$E_0 = \min_{\psi} \frac{\langle\psi|\hat{H}|\psi\rangle}{\langle\psi \mid \psi\rangle}$$

To obtain exact result, the minumum is taken over full Hilbert space, but we can obtain upper bound on ED by a fewparameter anstatz $|g_1, , g_2, \cdots\rangle$, e.g. $\langle r|g\rangle = e^{-\frac{1}{2}gr^2}$

$$E_0 \leq \frac{\langle g|\hat{H}|g\rangle}{\langle g|g\rangle}$$

The $|g\rangle$ obtained in this way is in this sense the "best" approximation to $|E_0\rangle$ in the Hilbert subspace $|\ \{g\}\rangle$.

We can do something similar in statistical physics.

Given T, H, we define the free-energy of an *arbitrary* distribution ρ by

$$\begin{aligned} F[\rho] &\equiv \sum \rho_\mu \left(H[\mu] + T\ln\left(\rho_\mu\right)\right) \\ &= \langle H\rangle_\rho - TS[\rho] \end{aligned}$$

where we've used the Gibbs entropy

$$S[\rho] = -\sum_{\mu} \rho_\mu \log(\rho_\mu)$$

Recall that $\rho_\beta = \frac{1}{Z_\beta} e^{-\beta H} = e^{-\beta(H-F_0)}$ minimizes F :

$$F[\rho] \geq F[\rho_\beta] \equiv F_0$$

(Because ρ_β maximizes the Gibbs entropy subject to $\langle E \rangle$)

Given a few-parameter ansatz $\rho_\mu(g)$, we can then define our best approximation as

$$\min_g F[\rho(g)] > F_0 \tag{12}$$

and use $\rho(g)$ to compute approximate observables.

Note that there is a simple way to see that (12)is true. Using $\rho_\beta = e^{-\beta(H-F_0)}$, we can write

$$\beta\left(F[\rho(g)] - F_0\right) = \beta\left[\langle H \rangle_{\rho(g)} - F_0\right] - S[\rho(g)]/k_B = \langle \ln(\rho(g)) \rangle_{\rho(g)} - \langle \ln(\rho(g))$$

showing that the difference between the approximate and true free energy is proportional to the “KL” divergence

$$S_{KL}(P \| Q) = \sum P_i \ln\left(\frac{P_i}{Q_i}\right) \geq 0$$

which is non-negative (see Shannon entropy lecture, eq. (2)).

It is convenient to parameterize $\rho(g)$ as a Boltzmann distribution corresponding to a fictitious Hamiltonian H_g,

$$\rho(g) \equiv \frac{1}{Z_g} e^{-\beta H_g} \ .$$

Consider, for example, the generalized Ising Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

where the notation $\langle i, j \rangle$ indicates that sites i and j are in contact (i.e. they are nearest neighbors) and each pair is only counted once. We might then choose $H_g = -g \sum_i \sigma_i$, where g is a parameter we will adjust. In terms of H_g, we have

$$F[\rho(g)] = \sum_\mu \frac{e^{-\beta H_g(\mu)}}{Z_g} \left(H(\mu) + \frac{1}{\beta} \ln \left(\frac{e^{-\beta H_g(\mu)}}{Z_g} \right) \right)$$

Defining $\quad e^{-\beta F_g} = Z_g = \sum_\mu e^{-\beta H_g(\mu)} \quad$ $(F_g \neq F[p(g)]!)$, we have

$$F[\rho(g)] = \langle H \rangle_g - \langle H_g \rangle_g + F_g$$

The lower bound

$$F[p(g)] - F_0 = F_g - F_0 + \langle H \rangle_g - \langle H_g \rangle_g \geq 0$$

is called the "Gibbs's inequality". It can be rewritten as

$$-\frac{1}{\beta} (\ln (Z_g) - \ln(Z)) + \langle H \rangle_g - \langle H_g \rangle_g \geqslant 0$$
$$\ln(Z) \geqslant \ln(Z_g) + \beta \left(\langle H_g \rangle_g - \langle H \rangle_g \right)$$

Let's try this for

$$H = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j,$$

$$H_g = -g \sum_i^N \sigma_i, \qquad \text{"Mean field"}$$

We need to compute $Z_g, \langle H_g \rangle_g, \langle H \rangle_g$

$$\begin{aligned}
Z_g &= \left(e^{\beta g} + e^{-\beta g}\right)^N = [2\cosh(\beta g)]^N \equiv z_g^N \\
F_g &= -\frac{1}{\beta} N \ln[2\cosh(\beta g)] \\
\langle H_g \rangle &= -g \left\langle \sum_i \sigma_i \right\rangle_g
\end{aligned}$$

which depends on the mean magnetization m_g under the fictitious Hamiltonian,

$$m_g \equiv \langle \sigma_i \rangle_g = N^{-1} \left\langle \sum_j \sigma_j \right\rangle_g = (Ng)^{-1} \partial_\beta \ln(Z_g) = \tanh(\beta g) \ .$$

And crucially, because under ρ_g the different spins are uncorrelated,

$$\begin{aligned}
&\rho_g(\sigma_1, \sigma_2, \ldots) = \rho_g(\sigma_1) \rho_g(\sigma_2) \ldots \\
&= \prod_i \underbrace{\left(\frac{e^{-\beta g \sigma_i}}{z_g} \right)}_{\rho_g(\sigma_i)}
\end{aligned}$$

we obtain for the ρ_g averaged energy

$$
\begin{aligned}
\langle H\rangle_g &= -J\sum_{\langle i,j\rangle}\langle\sigma_i\sigma_j\rangle_g = -J\sum_{\langle i,j\rangle}\langle\sigma_i\rangle_g\langle\sigma_j\rangle_g \\
&= -J\sum_{\langle i,j\rangle} m_g^2 = -J\sum_{\langle i,j\rangle}\tanh^2(\beta g) = \frac{N\zeta}{2}\tanh^2(\beta g)
\end{aligned}
$$

where ζ is the number of nearest neighbors each spin has - the so-called coordination number.

Putting it all together

$$
\begin{aligned}
&F[p(g)] = \langle H\rangle_g - \langle H_g\rangle_g + F_g \\
&= -J\frac{N\zeta}{2}m_g^2 + Ngm_g - \frac{1}{\beta}N\ln[2\cosh(\beta g)]
\end{aligned}
$$

Now, finally, we minimize ($m_g' = \partial_g m_g$):

$$
\begin{aligned}
&\partial_g F[\rho(g)] = N\partial_g\left[-J\frac{\zeta}{2}m_g^2 + gm_g - \frac{1}{\beta}\ln(\cosh(\beta g))\right] \\
&= N\left[-J\zeta m_g m_g' + gm_g' + m_g - m_g\right] = 0 \\
&g = J\zeta\cdot m_g = J\cdot\zeta\cdot\tanh(\beta g) \\
&\text{"self-consistency"}
\end{aligned}
$$

This has a simple physical interpretation: in $H = -J\sum_{\langle i,j\rangle}\sigma_i\sigma_j$, each σ_i sees "on average" a field $J\zeta\langle\sigma_i\rangle_g = J\zeta m_g$ induced by its neighors - suggesting $H \approx -J\zeta m_g\sum_i\sigma_i$. Since $H_g = -g\sum_i\sigma_i$, the condition is $g = J\zeta m_g$.

We can solve $g = J\zeta\tanh(\beta g)$ analytically for small g :

$$
g = J\zeta\beta g\left(1 - \frac{1}{3}(\beta g)^2\right) + \cdots
$$

Solution 1 : $g = 0. \longrightarrow m_g = 0$

But for $T\zeta\beta > 1$,

$$1 = J\zeta\beta\left(1 - \frac{1}{3}(\beta g)^2\right)$$

Solution 2: $\beta g = \pm\sqrt{3\left(1 - \frac{1}{J\zeta\beta}\right)}$

$$m_g = \pm\sqrt{3\left(1 - \frac{1}{J\zeta\beta}\right)}$$

For $J\zeta\beta > 1$, it can be verified this is lower-F solution: symmetry breaking!

Graphical Solution

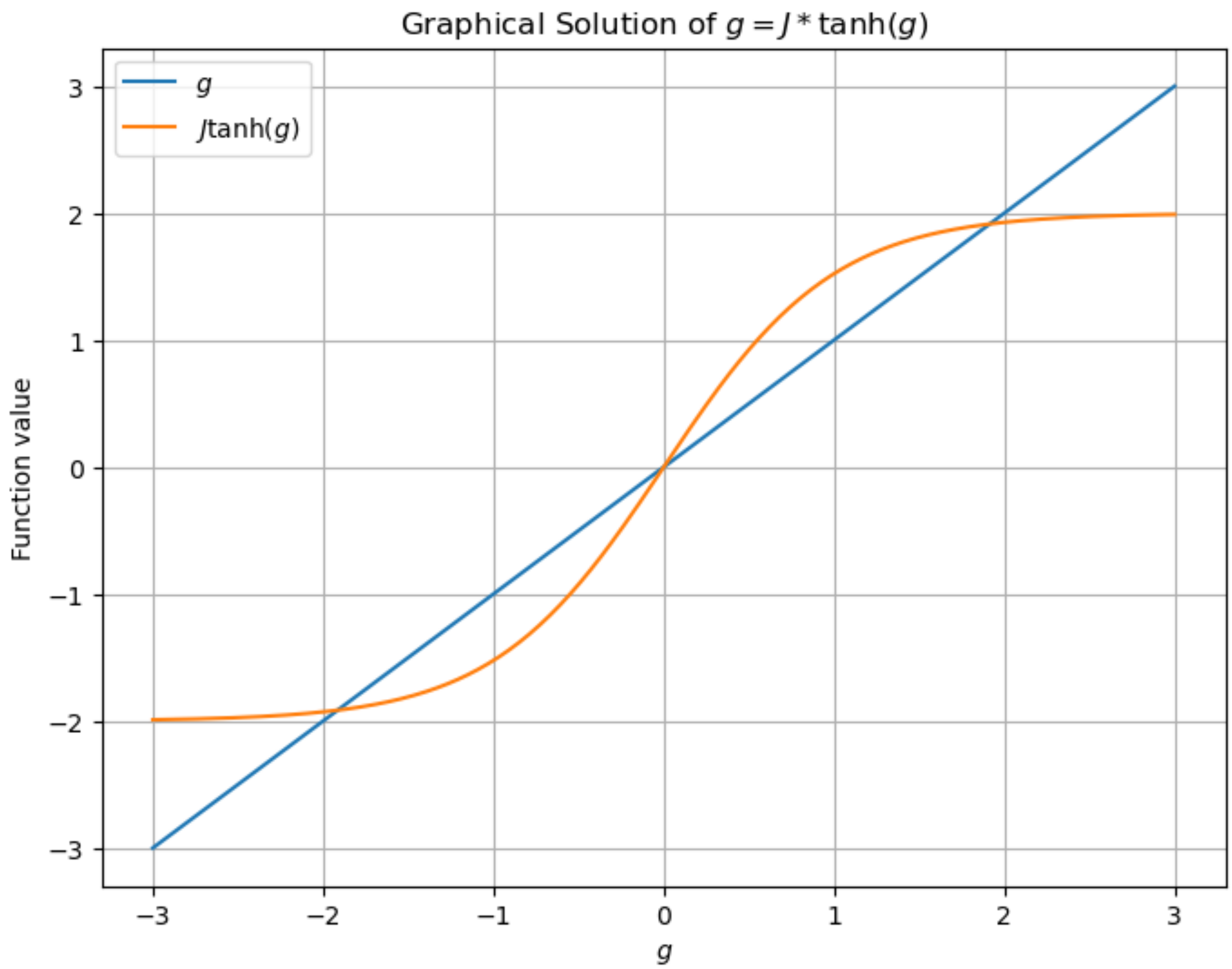
$$g = J\zeta\tanh(\beta g)$$

```
import numpy as np
import matplotlib.pyplot as plt

# Define the parameters
J = 2  # You can adjust J to see different behaviors
g_values = np.linspace(-3, 3, 400)  # Range of g values for plotting

# Define the functions
g_function = g_values
tanh_function = J * np.tanh(g_values)

# Plotting
plt.figure(figsize=(8, 6))
plt.plot(g_values, g_function, label='$g$')
plt.plot(g_values, tanh_function, label='$J \\tanh(g)$')
plt.title('Graphical Solution of $g = J * \\tanh(g)$')
plt.xlabel('$g$')
plt.ylabel('Function value')
plt.legend()
plt.grid(True)
plt.show()
```



Is this variational approximatition good? It knows about the lattice and dimensions $D = 1, 2, 3$ only through coordination number ζ. e.g., for square lattice $z = 2D$

But we know that the exact solution of 1D Ising model does not have symmetry breaking: the variational result is bad in 1D. On the other hand, if $(D, z) \to \infty$, you can verify $F[\rho(g)]$ is identical to the exact result of the all-to-all model: it is good in large D and large ζ.

For $D = 2, 3$, the accuracy is intermediate; it correctly predicts symmetry breaking but doesn't get T_c or $m \sim |T - T_c|^\beta$ quantitatively right.

Previous
‹ Finite dimensional Ising models