Statistical Physics of Computation - Written midterm

Vittorio Erba, Emanuele Troiani

Phys-512 – 2025/26

SCIPER:

Seat number:

Instructions:

- Duration of the midterm: 2 hours.
- The total number of points is 15.
- Allowed material: one A4 sheet recto-verso of notes (handwritten or printed) + common writing material.
- Write your answers in the provided spaces *in legible handwriting* and *in pen (no pencil).* If you need more space, you can use additional white pages. Please write in the provided space that an additional page was used, and on the additional page write clearly the number of the question you are answering. Write your SCIPER and seat number on all additional pages you use.
- Questions can be solved in any order. They *are not* ordered by difficulty or by topic.
- You can answer in English or French at you discretion.

Useful information:

- The Gaussian probability density function $\mathcal{N}(x;\mu;\sigma^2)$ with mean μ, variance σ^2 and evaluated at x is given by
$$\mathcal{N}(x;\mu;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \,. \tag{1}$$
- For $a > 0$ and any $b \in \mathbb{R}$ we have
$$\int \frac{e^{-ax^2/2+bx}}{\sqrt{2\pi}} \, \mathrm{d}x = \frac{e^{b^2/(2a)}}{\sqrt{a}} \,, \tag{2}$$
and
$$\int x \frac{e^{-ax^2/2+bx}}{\sqrt{2\pi}} \, \mathrm{d}x = \frac{b}{a} \frac{e^{b^2/(2a)}}{\sqrt{a}} \,. \tag{3}$$

Questions

1. (1pt) Evaluate $f(x)$, where $f(x)$ is defined as

$$f(x) = \lim_{n\to\infty} \frac{1}{n} \log \left[\int_{-\infty}^{\infty} e^{n\left(\frac{x}{3}t^3 - \frac{1}{4}t^4\right)} \, \mathrm{d}t \right] . \tag{4}$$

Solution: We use the saddle point approximation. We have

$$f(x) = \lim_{n\to\infty} \frac{1}{n} \log \left[\int_{-\infty}^{\infty} e^{n\left(\frac{x}{3}t^3 - \frac{1}{4}t^4\right)} \, \mathrm{d}t \right] = \max_{t\in\mathbb{R}} \left(\frac{x}{3}t^3 - \frac{1}{4}t^4 \right) = \max_{t\in\mathbb{R}} g(t)$$

the stationary points are given by

$$g'(t) = xt^2 - t^3 = 0 \iff t = 0 \text{ or } t = x$$

with values $g(0) = 0$ and $g(x) = x^4/12$. For $x \neq 0$. we see by the Taylor expansion of g at zero that the stationary point at $t = 0$ is a saddle (first non-zero term is cubic), while by computing the curvature

$$g''(t) = 2xt - 3t^2 \tag{5}$$

we see that the stationary point at $t = x$ is always a maximum. Thus for $x \neq 0$ we have $f(x) = x^4/12$. If $x = 0$ we have just a stationary point at zero, which is a maximum, giving $f(x = 0) = 0$ (notice that here one would need to expand g to fourth order to do the integral properly, but this alters only subleading terms). Thus we get $f(x) = x^4/12$.

2. (2.5pt) Consider the inference problem of retrieving a hidden vector $w^* \in \mathbb{R}^d$ from the observation of a dataset of n vectors $z^\mu \in \mathbb{R}^d$ with $\mu = 1, \dots, n$ and n associated labels $y^\mu \in \{-1, +1\}$. The hidden vector has components w_i^* ($i = 1, \dots, d$) that are independent Gaussian variables with mean zero and variance 1. The observed dataset is generated as follows. For each μ independently, y^μ equals $+1$ with probability $1/2$, and equals -1 with probability $1/2$. Then, z_i^μ (the i-th component of the vector z^μ, $i = 1, \dots, d$) is generated as a Gaussian variable with mean $y^\mu w_i^*$ and variance 1.
(1) Write the posterior distribution $P_{\text{post}}(w|\{z^\mu, y^\mu\}_{\mu=1}^n)$ as a function of the distributions of w^*, and $\{z^\mu, y^\mu\}_{\mu=1}^n$.
(2) Given a generic posterior distribution, how can you compute the Bayes-Optimal estimator with respect to the mean square error?
(3) Show that the Bayes-Optimal estimator $\hat{w}$ with respect to the mean square error of w^* is

$$\hat{w} = \frac{1}{n+1} \sum_{\mu=1}^{n} y^\mu z^\mu \,.$$

Solution: (1) The posterior on w is given by Bayes theorem by

$$P(w|\{x^\mu, y^\mu\}_{\mu=1}^n) \propto \prod_{\mu=1}^{n} P(x^\mu, y^\mu|w) P(w) \,.$$

We have by definition that

$$P(w) = \prod_{i=1}^{d} N(w_i; 0; 1) \tag{6}$$

and

$$P(x^\mu, y^\mu|w) = \frac{\delta_{y,1} + \delta_{y,-1}}{2} \prod_{i=1}^{d} N(x_i^\mu; y^\mu w_i; 1) \tag{7}$$

giving

$$P(w|\{x^\mu, y^\mu\}_{\mu=1}^n) \propto \prod_{i=1}^{d} \exp\left(-\frac{n+1}{2} w_i^2 + w_i \sum_{\mu=1}^{n} x_i^\mu y^\mu\right) \,, \tag{8}$$

where we discarded all terms independent on w.
(2,3) The BO estimator w.r.t. the MSE is given by the posterior mean. The posterior is factorized over dimension indices, and it's Gaussian, giving

$$\mathbb{E}_{w\sim\text{posterior}}[w_i] = \frac{1}{n+1} \sum_{\mu=1}^{n} x_i^\mu y^\mu \,. \tag{9}$$

3. (2pt) Consider a system with partition function $Z(\beta) = e^{\beta x^2}$ where x is a Gaussian variable with mean zero and variance 1 that we interpret as a disorder. Compute $\mathbb{E}_x[\log Z(\beta)]$ using the replica trick. Is the result correct?

Solution: First we compute $\mathbb{E}[Z^n]$

$$\mathbb{E}[Z^n] = \frac{1}{\sqrt{2\pi}} \int e^{-x^2/2+n\beta x^2}\, \mathrm{d}x = 1/\sqrt{1-2n\beta}$$

Next, we expand in small n

$$\mathbb{E}[Z^n] \approx 1 + n\beta$$

Finally.

$$\mathbb{E}[\log Z] = \lim_{n\to 0} \frac{\mathbb{E}[Z^n] - 1}{n} = \beta$$

The same quantity can be computed directly as

$$\mathbb{E}[\log Z] = \mathbb{E}[\beta x^2] = \beta \tag{10}$$

verifying the result.

4. (3pt) Recall that the Hamiltonian of the Curie-Weiss model with n spins is

$$H_n(s) = -\frac{1}{2n}\sum_{i,j=1}^{n} s_i s_j - h\sum_{i=1}^{n} s_i = -\frac{1}{2n}\Big(\sum_{i=1}^{n} s_i\Big)^2 - h\sum_{i=1}^{n} s_i\,. \tag{11}$$

where each $s_i \in \{-1, 1\}$. We found in class that free entropy function at inverse temperature β is

$$\phi(\beta, h) = \lim_{n\to\infty}\frac{1}{n}\log Z_n(\beta, h) = \max_m\Big\{-\frac{\beta}{2}m^2 + \log[2\cosh\big(\beta(h+m)\big)]\Big\}. \tag{12}$$

(1) What is the definition of $Z_n(\beta, h)$ in terms of $H_n(s)$?
(2) Re-derive the free entropy **without using delta functions**, using the Hubbard-Stratonovich transform

$$e^{\frac{a}{2}x^2} = \sqrt{\frac{a}{2\pi}}\int_{\mathbb{R}} e^{-\frac{az^2}{2}+zx}\,\mathrm{d}z, \qquad a > 0. \tag{13}$$

Addendum: here the provided form of the HS transform has a typo, with the correct formula being

$$e^{\frac{a}{2}x^2} = \sqrt{\frac{a}{2\pi}}\int_{\mathbb{R}} e^{-\frac{z^2}{2a}+zx}\,\mathrm{d}z, \qquad a > 0. \tag{14}$$

We did not count as error any incorrect statement that derived directly from this typo.

Solution: (1) The partition function is defined as

$$Z_n(\beta, h) = \sum_{s_1,\ldots,s_n=\pm 1} e^{-\beta H_n(s)}\,. \tag{15}$$

Taking $a = \frac{1}{n\beta}$ and $x = \beta\sum_i s_i$, we obtain from the HS transform

$$e^{\frac{\beta}{2n}(\sum_i s_i)^2} = e^{\frac{1}{2n\beta}(\beta\sum_i s_i)^2} = \sqrt{\frac{1}{2\pi n\beta}}\int_{\mathbb{R}}\exp\left[-\frac{n\beta}{2}m^2 + \beta m\sum_i s_i\right]\mathrm{d}m. \tag{16}$$

Hence

$$Z_n = \sqrt{\frac{1}{2\pi n\beta}}\int_{\mathbb{R}}\exp\Big(-\frac{n\beta}{2}m^2\Big)\sum_{s\in\{\pm1\}^n}\exp\left[(\beta h + \beta m)\sum_i s_i\right]\mathrm{d}m \tag{17}$$

$$= \sqrt{\frac{1}{2\pi n\beta}}\int_{\mathbb{R}}\exp\Big\{n\Big[-\frac{\beta}{2}m^2 + \log 2\cosh\big(\beta(h+m)\big)\Big]\Big\}\,\mathrm{d}m. \tag{18}$$

Finally, use the saddle point on m.

5. (2.5pt) Consider an inference problem in the Bayes-optimal setting, given as a function of a signal-to-noise ratio λ. Suppose that the BO estimator obtains a magnetization with the hidden signal m_{BO}, and that there exists an AMP algorithm that achieves a magnetization with the hidden signal m_{AMP} at convergence (when initialized from the prior). Assume that Nishimori's conditions hold also for AMP.
(1) What relation between m_{BO} and m_{AMP} makes you conclude that we have an algorithmically hard phase?
(2) What relation between m_{BO} and m_{AMP} makes you conclude that AMP reaches BO performance?
(3) Suppose that the state equation of the system is given by $F_\lambda(m) = 0$, and that the state evolution of AMP converges to a solution of $F_\lambda(m) = 0$. Under which condition on F_λ we can be sure a hard phase **does not** exists?

> **Solution:** (1) $m_{\mathrm{BO}} > m_{\mathrm{AMP}}$ strictly (but $m_{\mathrm{BO}} \neq m_{\mathrm{AMP}}$ was also counted correct).
> (2) $m_{\mathrm{BO}} = m_{\mathrm{AMP}}$.
> (3) If $F_\lambda(m) = 0$ has only one solution, then necessarily $m_{\mathrm{BO}} = m_{\mathrm{AMP}}$, as both magnetizations are solutions to the state equation.

6. (2pt) I have a Bayes Optimal problem. Can I assume replica symmetry? Under which additional assumption?

> **Solution:** Yes. By Nishimori (Bayes optimality) we know that the overlap distribution is the same as the magnetization distribution. If the magnetization distribution concentrates, then the overlap distribution concentrate, which is equivalent to replica symmetry.

7. (1pt) You are studying a system with an external parameter λ you are free to tune. You notice that by varying λ the order parameter m undergoes a phase transition. Draw two qualitative sketches of a m vs λ plot, one representing an example of a first-order (discontinuous), and the other representing an example of a second-order (continuous) phase transition.

> **Solution:** Case 1: any plot of $m(\lambda)$ with a jump discontinuity. Case 2: any plot of $m(\lambda)$ with an angular point.

8. (1pt) You are given a matrix Y from a spiked-Wigner model in dimension $N \gg 1$

$$Y_{ij} = \sqrt{\lambda/N}\, v_i v_j + G_{ij} \tag{19}$$

where G is a symmetric matrix with standard Gaussian entries and v is a vector that has standard Gaussian entries. Draw two qualitative sketches of the histogram of the eigenvalues, one for $\lambda \ll 1$ and the other for $\lambda \gg 1$.

> **Solution:** For small λ, the spectrum will be just the spectrum of G, which is a semicircle (any shape of the bulk here was counted as correct). For large λ the spectrum will be a similar bulk plus am outlying spike.