Institut for Theoretical Physics
University Leipzig

Prof. Dr. B. Rosenow
M. Kühn

Statistical Mechanics of Deep Learning - Problem set 5

Winter Term 2024/25

The problem set will be discussed in the seminar on **Monday 18.11.2024, 9:15**.

10. Volume of version space with spherical constraints *5+3 Points*

(a) Show that the surface area of the sphere, occupied by vectors $\boldsymbol{J}$ with the normalization constraint $\boldsymbol{J}^2 = N$, is for large N to leading order given by

$$\Omega = \int d^N\boldsymbol{J}\ \delta(\boldsymbol{J}^2 - N) \sim \exp\left(\frac{N}{2}[1 + \ln(2\pi)]\right).$$

Hint: Start by computing $A_N(r)$, the area of a sphere in N dimensions with radius r, using an N dimensional Gaussian integral, and then do an approximation for large N.

(b) Evaluate the volume of the version space defined by all vectors $\boldsymbol{J}$ that have an angle $\pi\varepsilon$ with a given direction defined by the vector $\boldsymbol{T}$, i.e. show that the following integral is for large N to leading order given by

$$\Omega_0(\varepsilon) = \int d\boldsymbol{J}\ \delta(\boldsymbol{J}^2 - N)\ \delta\left(\frac{\boldsymbol{J}\boldsymbol{T}}{N} - \cos(\pi\varepsilon)\right) \sim \exp\left(\frac{N}{2}\left[1 + \ln(2\pi) + \ln\sin^2(\pi\varepsilon)\right]\right).$$

Hint: The surface area of a spherical cap up to an angle ϕ can be calculated as $A_N^{cap}(r) = \int_0^\phi A_{N-1}(r\sin\theta)rd\theta$.

11. Gaussian joint probability density function *4+4 Points*

Consider the auxiliary variables

$$\lambda_\mu = \frac{1}{\sqrt{N}}\boldsymbol{J}\boldsymbol{\xi}^\mu\ ,\ u_\mu = \frac{1}{\sqrt{N}}\boldsymbol{T}\boldsymbol{\xi}^\mu,$$

where $\boldsymbol{J}$ and $\boldsymbol{T}$ are the student and teacher vectors, while $\boldsymbol{\xi}^\mu \in \mathbb{R}^N$ are the example vectors with components ξ_i^μ drawn from the distribution

$$P(\boldsymbol{\xi}) = \prod_j^N \left[\frac{1}{2}\delta(\xi_j + 1) + \frac{1}{2}\delta(\xi_j - 1)\right]$$

.

(a) Show that the joint probability density $P(\lambda, u)$ is indeed a Gaussian probability density . Start from

$$P(\lambda, u) = \left\langle \left\langle \delta\left(\lambda - \frac{1}{\sqrt{N}}\boldsymbol{J}\boldsymbol{\xi}\right) \delta\left(u - \frac{1}{\sqrt{N}}\boldsymbol{T}\boldsymbol{\xi}\right)\right\rangle\right\rangle_{\boldsymbol{\xi}}$$

The average in $P(\lambda, u)$ is with respect to a randomly chosen example $\boldsymbol{\xi}$.
Hint : you may use the Hubbard-Stratonovich transformation

$$\int Dt\ e^{bt} = e^{b^2/2}$$

$$\text{where } Dt := \frac{dt}{\sqrt{2\pi}}\exp(-t^2/2)$$

(b) Show that the distribution $P(\lambda, u)$ has the moments

$$\langle\langle\lambda\rangle\rangle = \langle\langle u\rangle\rangle = 0$$
$$\langle\langle\lambda^2\rangle\rangle = \langle\langle u^2\rangle\rangle = 1$$
$$\langle\langle\lambda u\rangle\rangle = \frac{\boldsymbol{J}\boldsymbol{T}}{N} = R.$$