Institut for Theoretical Physics
University Leipzig

Prof. Dr. B. Rosenow
A. Afanah , M. Staats

Statistical Mechanics of Deep Learning - Problem set 9

Winter Term 2023/24

Hand in: Friday, 15.12 at 10:00 am, you can upload your solutions to the course webpage on Moodle platform.

18. On-line learning of the perceptron rule *9+3 Points*

In this problem set, we aim to derive the dynamical equations of the order parameters ρ and Q of the perceptron rule,

$$\frac{d\rho}{d\alpha} = \langle F\, u \rangle \tag{1}$$

$$\frac{dQ}{d\alpha} = \langle F(F + 2t) \rangle \tag{2}$$

with F here denoting the learning amplitude function of the perceptron algorithm given by

$$F = \eta\, \theta(-tu)\, \mathrm{sgn}(u).$$

The averages in eqs.(1) and (2) are over the correlated gaussian random variables u and $t/\sqrt{Q}$ with zero mean and the second moments

$$\langle u^2 \rangle = 1,\ \left\langle \frac{t^2}{Q} \right\rangle = 1,\ \left\langle ut/\sqrt{Q} \right\rangle = \rho/\sqrt{Q} = R$$

(a) Show that

$$\langle \theta(-tu) \rangle = \frac{1}{\pi}\ \arccos\left(\frac{\rho}{\sqrt{Q}}\right)$$

$$\langle t\, \theta(-tu)\, \mathrm{sgn}(u) \rangle = \frac{\rho - \sqrt{Q}}{\sqrt{2\pi}}$$

$$\langle \theta(-tu)\, |u| \rangle = \frac{\sqrt{Q} - \rho}{\sqrt{2\pi\, Q}}$$

(b) Use the results obtained in (a) to derive the order parameters equations

$$\frac{d\rho}{d\alpha} = \frac{\eta}{\sqrt{2\pi}}\ \left(1 - \frac{\rho}{\sqrt{Q}}\right)$$

$$\frac{dQ}{d\alpha} = \frac{\eta^2}{\pi}\ \arccos\left(\frac{\rho}{\sqrt{Q}}\right) + \sqrt{\frac{2}{\pi}}\ \eta\ (\rho - \sqrt{Q})$$