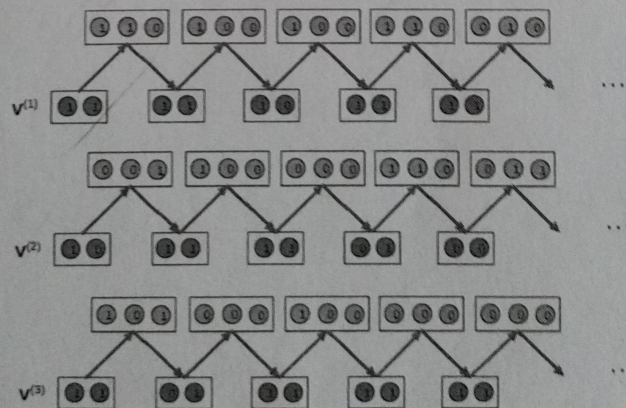


## Duboko učenje

pismeni ispit

1. Razmatramo sigmoidalnu transformaciju podataka u izlaznom sloju dubokog modela kojeg učimo optimiranjem i) kvadratnog gubitka ii) negativne log-izglednosti. Izvedite gradijente obaju gubitaka s obzirom na ulaze sigmoidalne funkcije. Koji od dvaju gubitaka je prikladniji za slučaj klasifikacije? Zašto? Napišite kod za računanje gradijenata pod numpyjem pod pretpostavkom da podatci dolaze u grupama proizvoljne veličine.
2. Razmatramo učenje višerazredne logističke regresije optimiranjem unakrsne entropije stohastičkim gradijentnim spustom uz zadanu stopu učenja  $\delta = 1$ . Zadan je sljedeći skup za učenje:  $x_1 = (1, 1, 1, 0)$ ,  $y_1 = [1, 0]$ ,  $x_2 = (0, 1, 1, 1)$ ,  $y_2 = [0, 1]$ . Početno stanje matrice težina i vektora pomaka zadano je s  $\mathbf{W}^0 = \begin{bmatrix} 0.5 & -0.5 & -0.5 & 0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 \end{bmatrix}$ ,  $\mathbf{b}^0 = \begin{bmatrix} 0 & 0 \end{bmatrix}$ .
  - (a) Odredite jednadžbe gradijenata svih parametara.
  - (b) Provedite učenje podatkom  $(x_1, y_1)$ , bez korištenja ikakve regularizacije. Izračunajte gradijente i nove vrijednosti parametara.
  - (c) Provedite jednu epohu učenja počevši od  $\mathbf{W}^0$  i  $\mathbf{b}^0$ , uz regularizaciju ispuštanjem (engl. dropout) ulaznih značajki s vjerojatnošću  $p(\mu_{x_i}) = 0.5$ ,  $i = 1, 2, 3, 4$ . Pretpostavite sljedeći raspored ispuštanja:
    - prva mini-grupa je  $(x_1, y_1)$ , ugašeni su 1. i 3. ulaz,
    - druga mini-grupa je  $(x_2, y_2)$ , ugašeni su 2. i 4. ulaz.
 Izračunajte gradijente i nove vrijednosti parametara.
  - (d) Izračunajte izlaz naučene mreže za ulaze  $x_a = (0, 1, 0, 1)$  i  $x_b = (1, 0, 1, 0)$ .
3. Razmatramo ograničeni Boltzmannov stroj s dva elementa u vidljivom i tri elementa u skrivenom sloju. Gibbsovo uzorkovanje za mini-grupu od tri uzorka  $v^{(1)}$ ,  $v^{(2)}$  i  $v^{(3)}$  prikazano je na slici. Potrebno je izračunati korekciju i) za težinu  $w_{13}$  koja povezuje elemente  $v^{(1)}$  i  $h^{(3)}$  te ii) za težinu  $w_{23}$  koja povezuje  $v^{(2)}$  i  $h^{(3)}$ . U rješavanju koristite algoritam CD-3 i koeficijent učenja  $\eta=0.01$ .





4. Za izlazni sloj varijacijskog autoenkodera odabrali smo Rayleighovu distribuciju:

$$p(x|\sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}$$

Izvedite izraz za drugu komponentu funkcije cilja varijacijskog autoenkodera:

$$E_{q_{\varphi}(\mathbf{z}|\mathbf{x}^{(i)})} [\log(p_{\theta}(\mathbf{x}^{(i)}|\mathbf{z})]$$

Pretpostavite da je dovoljno samo jednom uzorkovati skriveni sloj jer su mini-grupe dovoljno velike. Skicirajte odgovarajući dekođer.

5. Učiteljica Svjetlana suočena je s teškim problemom - učenici njenog 2.D razreda iznimno su nestašni. Naime, kad stanu u red za prebrojavanje koliko ih je prisutno na nastavi, neki od njih se nakon što su prebrojani vrte na kraj reda, te ih zbog svoje zaboravljivosti Svjetlana ponovno prebroji, i ponekad po kraju (dugotrajnog) brojanja izgleda kao da je mnogo više djece na nastavi nego što ih ima u razredu - što je stvarno nemoguće. Svjetlana je u novinama pročitala o novoj metodi povratnih neuronskih mreža, te ih je odlučila iskoristiti za svoj problem.

Zadatak je postavila kao u nastavku: svaki indeks skrivenog stanja povratne neuronske mreže odgovara jednom učeniku, te na tom indeksu želi imati informaciju je li se učenik pojavio taj dan ili nije (0 ako nije, 1 ako je). Osim toga, kako bi skratila vrijeme brojanja, želi da povratna mreža na izlazu ispiše (1 - broj učenika koji nedostaju). Dakle, u slučaju da su svi učenici prisutni na izlazu je 1, ako fali jedan učenik, na izlazu je 0 a ako fali dvoje učenika, na izlazu je -1 itd.

U Svjetlaninom razredu je troje učenika, te je svakom od njih dodijeljen one-hot kod. Na ulaz mreže u slijedu dolaze one-hot kodovi učenika. Radi jednostavnosti, prijenosna funkcija na izlazu je identitet (umjesto funkcije softmax), a dimenzionalnost izlaza je 1. Skriveno stanje mreže je inicijalizirano na vektor nula. Ukoliko je poznato da je matrica  $\mathbf{U}$  jedinična, matrica  $\mathbf{W}$  dijagonalna, odredite vektore pristranosti  $\mathbf{b}$  i  $\mathbf{c}$ , matrice težina  $\mathbf{W}$  i  $\mathbf{V}$  te prijenosnu funkciju skrivenog sloja  $f$ .