



یادگیری عمیق

پاییز ۱۴۰۳
استاد: دکتر فاطمی زاده

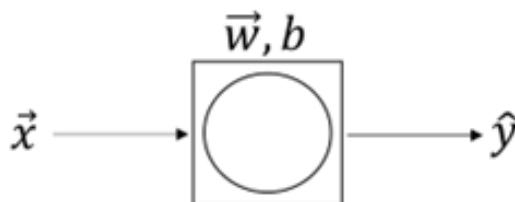
گردآورندگان: کیان کمندی، امیررضا ولایی

تمرین دوم شبکه های پرسپترون، رگولاسیون و بهینه سازی مهلت ارسال: ۱۸ آبان

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- در طول ترم امکان ارسال با تاخیر پاسخ همه ی تمارین تا سقف ۵ روز و در مجموع ۱۵ روز، وجود دارد. پس از گذشت این مدت، پاسخ های ارسال شده پذیرفته نخواهند بود. همچنین، به ازای هر روز تأخیر غیر مجاز ۱۰ درصد از نمره تمرین به صورت ساعتی کسر خواهد شد.
- همکاری و همفکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد. (دقت کنید در صورت تشخیص مشابهت غیرعادی برخورد جدی صورت خواهد گرفت.)
- در صورت همفکری و یا استفاده از هر منابع خارج درسی، نام همفکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.
- نتایج و پاسخ های خود را در یک فایل با فرمت zip به نام HW۲-Name-StudentNumber در سایت [Quera](#) قرار دهید. برای بخش عملی تمرین نیز لینک گیت هاب که تمرین و نتایج را در آن آپلود کرده اید قرار بدهید. دقت کنید هر سه فایل نوتبوک تکمیل شده بخش عملی را در گیت هاب قرار دهید.
- لطفا تمامی سوالات خود را از طریق کوثرای درس مطرح بکنید (برای اینکه تمامی دانشجویان به پاسخ های مطرح شده به سوالات دسترسی داشته باشند و جلوی سوالات تکراری گرفته شود، به سوالات در بسترهای دیگر پاسخ داده نخواهد شد).
- دقت کنید کدهای شما باید قابلیت اجرای دوباره داشته باشند، در صورت دادن خطا هنگام اجرای کدتان، حتی اگر خطا بدلیل اشتباه تایپی باشد، نمره صفر به آن بخش تعلق خواهد گرفت.

سوالات نظری (۳۰۰ نمره)

۱. (۴۰ نمره) فرض کنید یک مدل تک لایه و تنها با یک نورون مطابق شکل زیر داریم که دارای تابع فعالیت sigmoid می باشد.



برای یک مسئله طبقه بندی باینری تابع هزینه زیر را در نظر می گیریم:

$$E_{w,b} = - \sum_n y_n \ln \hat{y}(x_n) + (1 - y_n) \ln(1 - \hat{y}(x_n))$$

نشان دهید تابع هزینه دارای مینیمم است. سپس یک رابطه بازگشتی برای رسیدن به نقطه مینیمم بنویسید.

۲. (۵۰ نمره)

به سوالات زیر در مورد Batch Normalization (BN) پاسخ دهید.

(آ) مشکل covariate shift را در شبکه های عصبی بیان کرده و توضیح دهید BN چگونه آنرا حل می کند.

(ب) توضیح دهید BN چگونه به generalization شبکه کمک می کند.

(پ) یک BN ساده شده را در نظر بگیرید که داده ها را بر انحراف معیار تقسیم نمی کنیم. یعنی صرفاً برای ورودی های $[x_1, x_2, \dots, x_n]$ داریم: $\hat{x}_i = x_i - \mu$ که در آن $\mu = \frac{1}{n} \sum_{j=1}^n x_j$ می باشد (n تعداد داده ها در mini-batch است). خروجی های ما نیز $[y_1, y_2, \dots, y_n]$ هستند که در آن $y_i = \gamma \hat{x}_i + \beta$ می باشد. فرض کنید در انتهای شبکه عمیق یک تابع هزینه L تعریف شده باشد. $\frac{\partial L}{\partial x_i}$ را بر حسب $\frac{\partial L}{\partial y_j}$ برای $j = 1, \dots, n$ محاسبه کنید.

(ت) در قسمت پ برای دو حالت $n = 1$ و $n \rightarrow \infty$ مقدار $\frac{\partial L}{\partial x_i}$ را بدست آورید. چه نتیجه ای می گیرید؟

۳. (۵۰ نمره) یک شبکه عصبی دو لایه برای طبقه بندی K کلاس را که روابط آن بصورت زیر است را در نظر بگیرید. در این روابط x ورودی مدل به ابعاد $1 \times d_x$ و $y \in \{0, 1\}^K$ لیبیل های انکود شده بصورت one-hot می باشند. همچنین لایه مخفی دارای d_a نورون است.

$$z^{(1)} = W^{(1)}x + b^{(1)}$$

$$\hat{a}^{(1)} = \text{LeakyReLU}(z^{(1)}, \alpha = 0.01)$$

$$a^{(1)} = \text{Dropout}(\hat{a}^{(1)}, p = 0.2)$$

$$z^{(2)} = W^{(2)}a^{(1)} + b^{(2)}$$

$$\hat{y} = \text{softmax}(z^{(2)})$$

$$L = \sum_{i=1}^K -y_i \log(\hat{y}_i)$$

(آ) $\frac{\partial \hat{y}_k}{\partial z_i^{(2)}}$ را محاسبه کنید و پاسخ خود را بر حسب \hat{y} ساده کنید.

(ب) فرض کنید مقدار k -ام بردار لیبیل y برابر با ۱ و بقیه مقادیر آن صفر باشد) $\frac{\partial L}{\partial z_i^{(2)}}$ ($y_k = 1$ و $y_i = 0; i \neq k$ را محاسبه کرده و جواب خود را بر حسب \hat{y} ساده کنید.

(پ) $\frac{\partial L}{\partial W^{(1)}}$ را محاسبه کنید.

۴. (۴۰ نمره) نشان دهید ماتریس Hessian یک تبدیل مثل $\psi(u, v, z) = y(u, v, z)$ را می توان به صورت ماتریس ژاکوبی گرادیان این تبدیل نوشت. توجه کنید که متغیرهای u, v, z نک بعدی و y نیز تابعی بر حسب آنها است.

۵. (۴۰ نمره) تابع خطا در یک شبکه با اعمال **Dropout** گوسی ضربی به شرح زیر است.

$$J_1 = 0.5 \left(y_d - \sum_{k=1}^n \delta_k W_k x_k \right)^2$$

که در آن توزیع $\delta_k \sim Normal(1, \sigma^2)$ می باشد. مقدار امید ریاضی گرادیان تابع هدف نسبت به متغیر W_i را نوشته و تا حد ممکن ساده کنید.

آیا می توانید تعبیری از Regularization با استفاده از این نوع Dropout ارایه دهید؟ در صورت امکان تابع هدف Regularized-Non را هم معرفی کنید.

۶. (۴۰ نمره) همگرایی روش نیوتون را برای تابع $f(x) = g'(x)$ اثبات کنید. فرض کنید که f مشتق درجه دوم پیوسته دارد و در نقطه بهینه x^* مقدار $f(x^*) = 0$ و $f'(x^*) \neq 0$ دارد. سپس نشان دهید که روش مرتبه دوم نیوتون برای پیدا کردن x^* که نقطه بهینه تابع g است نیز همگراست.

۷. (۴۰ نمره) یک شبکه عصبی برای طبقه بندی چندکلاسی با K کلاس را در نظر بگیرید که لایه خروجی آن از تابع softmax استفاده می کند. ورودی به لایه softmax به صورت $\mathbf{z} = [z_1, z_2, \dots, z_K]^T \in \mathbb{R}^K$ است و خروجی های softmax به صورت زیر تعریف می شوند:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}, \quad i = 1, 2, \dots, K.$$

برچسب های واقعی کلاس ها با یک بردار one-hot $\mathbf{y} = [y_1, y_2, \dots, y_K]^T$ نمایش داده می شوند، که در آن $y_k = 1$ اگر کلاس k کلاس صحیح باشد، و در غیر این صورت $y_k = 0$. تابع زیان cross-entropy به صورت زیر تعریف می شود:

$$L(\mathbf{z}, \mathbf{y}) = - \sum_{k=1}^K y_k \log \hat{y}_k$$

(آ) اثبات کنید که گرادیان تابع زیان نسبت به \mathbf{z} به صورت زیر است:

$$\nabla_{\mathbf{z}} L = \hat{\mathbf{y}} - \mathbf{y}.$$

(ب) ماتریس هسین و مثبت نیمه معین بودن:

(الف) ماتریس هسین $H \in \mathbb{R}^{K \times K}$ تابع زیان نسبت به \mathbf{z} را محاسبه کنید.

(ب) اثبات کنید که H مثبت نیمه معین است.

(ج) با استفاده از نتیجه قسمت ۲، نتیجه گیری کنید که آیا تابع زیان آنتروپی متقاطع $L(\mathbf{z}, \mathbf{y})$ یک تابع محدب نسبت به \mathbf{z} است یا خیر.

سوالات عملی (۳۰۰ نمره)

۱. (۱۵۰ نمره) فایل های مربوط به این تمرین در پوشه hw2_1 قرار داده شده است. در این تمرین شما باید قسمت های خواسته شده در هر چهار فایل قرار داده شده یعنی solver.py, layers.py, optimizer.py و hw2_1.ipynb را کامل کنید. در فایل نوتبوک hw2_1.ipynb توضیحات مربوطه داده شده است.

۲. (۱۵۰ نمره) فایل های مربوط به این تمرین در پوشه hw2_2 قرار داده شده است. با توجه به توضیحات موجود در hw2_2.ipynb این نوتبوک را کامل کرده و به سوالات داخل آن پاسخ دهید.