باسمه تعالی

دانشگاه صنعتی شریف

دانشکده مهندسی برق

مقدمه ای بر یادگیری ماشین - گروه دکتر امینی
پاییز ۱۴۰۱

تمرین سری سوم

---

۱. مهلت تحویل این تمرین مطابق تاریخ اعلام شده در سامانه CW می باشد.

۲. ۱۴ روز تاخیر مجاز برای تحویل تمارین در اختیار شما خواهد بود.

۳. سقف برای تحویل هر تمرین ۷ روز خواهد بود و پس از آن پاسخنامه تمرین منتشرخواهد شد.

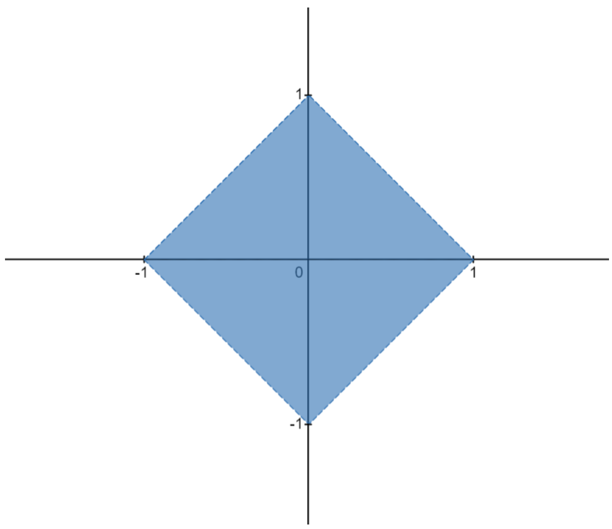
۴. ابهامات و اشکالات خود را درباره تمرین تئوری با آقای رامی و درباره تمرین عملی با آقای پورغنی مطرح کنید.

@mamin_rami
@Amin_pourghani

---

۱ شبکه عصبی با دو لایه مخفی

یک شبکه عصبی در نظر بگیرید که ورودی آن $\mathbf{x} \in R^2$ است. این شبکه عصبی دو لایه مخفی(با احتساب لایه خروجی) دارد. تابع فعالسازی این شبکه، تابع sign می باشد. وزن ها و بایاس های این شبکه رو به گونه ای تعیین کنید که در ناحیه آبی رنگ، خروجی شبکه ۱ ودر باقی نواحی ۱- باشد.



۲ مشتق برداری

توابع زیر را درنظر بگیرید:

$$\mathbf{f_1}(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} \frac{1}{\pi}x_1 \sin(\pi x_2) \\ e^{x_1-1}x_2^2 \\ x_1x_2 \end{bmatrix}$$

$$\mathbf{f_2}(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) = \begin{bmatrix} x_1+x_2+x_3 \\ x_1^2+x_2^2+x_3^2 \end{bmatrix}$$

$$\mathbf{f}(\mathbf{x}) = (\mathbf{f_2} \circ \mathbf{f_1})(\mathbf{x})$$

مطلوب است محاسبه مشتق زیر در نقطه $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$:

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$$

۳ الگوریتم Backpropagation

یک شبکه عصبی با مشخصات زیر در نظر بگیرید:

$\mathbf{x}$: بردار ورودی

$\mathbf{z_1} = \mathbf{W_1x}$: خروجی لایه اول، قبل از اعمال لایه فعالسازی

$\mathbf{y_1} = f(\mathbf{z_1})$: خروجی لایه اول

z2 is scalar

$\mathbf{z_2} = \mathbf{w_2^T y_1}$: خروجی شبکه قبل از اعمال تابع فعالسازی

خروجی شبکه : $y = f(\mathbf{z_2})$

همچنین داریم : $\mathbf{x} \in R^n$ ،$\mathbf{W_1} \in R^{m_1 \times n}$ ،$\mathbf{w_2} \in R^{m_1}$ ،$y \in R$

تابع f نیز یک تابع فعالسازی دلخواه است. مطلوب است محاسبه مشتقات زیر:

$$\frac{\partial y}{\partial \mathbf{W_1}}, \frac{\partial y}{\partial \mathbf{w_2}}$$

۴ کرنل گوسی

در این مسئله میخواهیم بررسی کنیم که رابطه معروف کرنل گوسی، چگونه به دست می آید. فرض کنید $x \in R$ باشد.

نگاشت $\mathbf{\Phi}(x)$، x را از R به فضای بینهایت بعدی می برد به اینصورت که $\mathbf{\Phi}(x)$ یک بردار با بینهایت المان است که المان n ام آن برابر مقدار زیر است:

$$\Phi_n(x) = \frac{1}{\sqrt{n!}} x^n e^{-\frac{x^2}{2}}$$

$$0 \leq n \leq \infty$$

حال تعریف میکنیم:

$$K(x, x') =< \mathbf{\Phi}(x), \mathbf{\Phi}(x') >$$

نشان دهید:

$$K(x, x') = e^{-\frac{|x-x'|^2}{2}}$$

۵ مدل SVM

یک دیتاست با دونقطه با ویژگی های یک بعدی در نظر بگیرید:

$$(x_1 = 0, y_1 = -1), (x_2 = \sqrt{2}, y_2 = 1)$$

یک نگاشت به صورت زیر در نظر بگیرید که ویژگی هارا از فضای یک بعدی به فضای سه بعدی می برد:

$$\mathbf{\Phi}(x) = [1, \sqrt{2}x, x^2]^T$$

بردار بهینه را به صورت زیر تعریف می کنیم:

$$min||\mathbf{w}||^{\mathbf{2}} \quad s.t:$$

$$y_1(\mathbf{w^T}\mathbf{\Phi}(x_1) + w_0) \geq 1$$

$$y_2(\mathbf{w^T}\mathbf{\Phi}(x_2) + w_0) \geq 1$$

الف) یک بردار پیدا کنید که با بردار بهینه $\mathbf{w}$ موازی باشد. (راهنمایی: برای پیدا کرد یک بردار موازی با بردار بهینه، به این نکته توجه کنید که بردار بهینه و هر بردار موازی با آن، بر مرز تصمیم گیری عمود است)

ب) چه مارجینی توسط بردار بردار بهینه به دست خواهد آمد؟ (راهنمایی: توجه کنید که بردار بهینه، یک برداری است که دو نقطه را توسط یک خط (که مرز تصمیم گیری نام دارد) با بیشترین فاصله از هم جدا میکند.مارجین نیز فاصله هر نقطه تا مرز تصمیم گیری است که طبیعتا برای هر دونقطه یکسان است)

ج) حال مقدار مارجین را میدانیم. همچنین میدانیم که رابطه زیر برقرار است:

$$margin = \frac{1}{||\mathbf{w}||}$$

پس میتوانیم از رابطه بالا، مقدار $||\mathbf{w}||$ را محاسبه کنیم. راستای $\mathbf{w}$ نیز در بخش ب به دست آوردیم. باتوجه به این نکات، بردارد $\mathbf{w}$ را به دست آورید.

د) حال مقدار w_0 را پیدا میکنیم. چون هردو نقطه روی مرز تصمیم گیری هستند، نامساوی های مسئله بهینه سازی به تساوی تبدیل می شوند. با توجه به این موضوع، مقدار w_0 را بیابید.

ه) حال تابع جدا ساز که به صورت زیر تعریف میشود را به صورت صریح بیابید:

$$f(x) = w_0 + \mathbf{w^T}\mathbf{\Phi}(\mathbf{x})$$