تمرین ۱. فرض کنید $\mathcal{X}$ مکعب دودویی $\{۰,۱\}^n$ باشد. برای مجموعه‌ی $I \subset \{۱,۲,\ldots,n\}$ تابع زوجیت h_I را به این شکل تعریف می‌کنیم:
برای بردار $X = (x_۱,\ldots,x_n) \in \{۰,۱\}^n$ داریم

$$h_I(X) = (\sum_{i\in I} x_i) \mod ۲$$

بعد VC برای کلاس تمام توابع زوجیت $\mathcal{H}_{n-parity} = \{h_I : I \subset \{۱,۲,\ldots,n\}\}$ چقدر است؟

پاسخ ۱. با توجه به این‌که $|\mathcal{X}| = ۲^n$ بنابراین داریم:

$$\text{VCdim}(\mathcal{H}) \leq \log|\mathcal{X}| = n$$

حال بردارهای استاندارد $e_۱,\ldots,e_n$ را در نظر بگیرید. این مجموعه توسط $\mathcal{H}$ خرد می‌شود. زیرا برای هر بردار، کافی است زیرمجموعه‌ای از $\{۱,\ldots,n\}$ را در نظر بگیریم که تنها شامل اعضای ۱ بردار فوق است. مثلا برای بردار $(۱,۱,۰,\ldots,۰)$ زیرمجموعه‌ی $I = \{۱,۲\}$ را در نظر می‌گیریم.

تمرین ۲. فرض کنید $\mathcal{H}$ کلاس بازه‌های علامت‌دار باشد به عبارتی $\mathcal{H} = \{h_{a,b,s} : a \leq b, s \in \{-۱,۱\}\}$ به طوری که:

$$h_{a,b,s}(x) = \begin{cases} s & x \in [a,b] \\ -s & x \notin [a,b] \end{cases}$$

در این صورت $\text{VCdim}(\mathcal{H})$ را حساب کنید.

پاسخ ۲. نشان می‌دهیم که $\mathcal{H}$ هر سه نقطه‌ای را خرد می‌کند ولی هیچ چهار نقطه‌ای را خرد نمی‌کند. فرض کنید سه نقطه‌ی $c_۱ < c_۲ < c_۳$ داده شده است. حال بازه‌های $[c_۱,c_۳]$، $[c_۱,c_۲]$، $[c_۲,c_۳]$ و $[a_۱,a_۲]$ که در آن $c_۱ < a_۱ < c_۲ < a_۲ < c_۳$ توابع دلخواه را به دست می‌دهد (با انتخاب‌های مختلف ۱ و ۱- روی بازه). حال چهار نقطه‌ی دلخواه $c_۱ < c_۲ < c_۳ < c_۴$ را در نظر بگیرید. با این چهار نقطه نمی‌توان حالت $(۱,-۱,۱,-۱)$ را ساخت. زیرا برای این‌که $۱,-۱,۱$ به وجود آید باید -۱ روی یک بازه بین $c_۱$ و $c_۳$ تعریف شود و بنابراین مولفه‌ی چهارم باید ۱ باشد و نمی‌تواند -۱ باشد. بنابراین $\text{VCdim}(\mathcal{H}) = ۳$.

تمرین ۳. نشان دهید مساله‌ی *ERM* برای رگرسیون خطی نیست به تابع هزینه $l(h,(x,y)) = |h(x) - y|$ یک برنامه‌ریزی خطی است؛ به عبارتی نشان دهید که چطور می‌توان مساله‌ی

$$\min_w \sum_{i=۱}^{m} |<w,x_i> - y_i|$$

را به شکل یک برنامه‌ریزی خطی نوشت.

پاسخ ۳. با توجه به تعریف قدرمطلق برای هر عدد حقیقی $|c| \leq a$ داریم $-a \leq c \leq a$ بنابراین معادله‌ی فوق را می‌توان به شکل زیر بیان کرد:

$$\forall ۱ \leq i \leq m, -a_i \leq <w,x_i> - y_i \leq a_i$$

$$\min \sum_i a_i$$

تمرین ۴. فرض کنید الگوریتم پرسپترون را به این شکل تغییر داده‌ایم: در هر مرحله به جای $w^{(t+۱)} = w^{(t)} + y_i x_i$ قرار می‌دهیم $w^{(t+۱)} = w^{(t)} + \eta y_i x_i$ که در آن $\eta > ۰$. ثابت کنید تعداد مراحل لازم در الگورتیم جدید با الگوریتم اصلی برابر است.

پاسخ ۴. کافی است مراحل اثبات در الگوریتم اصلی را برای الگوریتم جدید اجرا کنیم:

$$
\begin{aligned}
< w^*, w^{(t+۱)} > - < w^*, w^{(t)} > &= \eta y_i < w^*, x_i > \geq \eta \\
< w^*, w^{(T+۱)} > &= \sum_{i=۱}^{T} (< w^*, w^{(t+۱)} > - < w^*, w^{(t)} > \geq \eta T \\
||w^{(t+۱)}||^۲ &= ||w^{(t)} + \eta y_i x_i||^۲ \\
&= ||w^{(t)}||^۲ + ۲\eta y_i < w^{(t)}, x_i > + \eta^۲ y_i^۲ ||x_i||^۲ \\
&\leq ||w^{(t)}||^۲ + \eta^۲ R^۲
\end{aligned}
$$

$$
\Rightarrow ||w^{(T+۱)}||^۲ \leq \eta^۲ T R^۲ \Rightarrow ||w^{(T+۱)}|| \leq \sqrt{T} \eta R
$$

$$
\begin{gathered}
۱ \geq \frac{< w^*, w^{(T+۱)} >}{||w^*|| ||w^{(T+۱)}||} \geq \frac{\eta T}{B \sqrt{T} \eta R} = \frac{\sqrt{T}}{BR} \\
\Rightarrow T \leq B^۲ R^۲
\end{gathered}
$$