4.1

یک دنباله‌ی کمکی از متغیرها به صورت $(s_1,\dots,s_m)$ تعریف می‌کنیم.

با توجه به راهنمائی مسئله، ERM معادل با مینیمم کردن رابطه خطی $\sum_{i=1}^{m} s_i$

با توجه به محدودیت های زیر:

$$(\forall i\in[m])\quad \langle w, X_i\rangle - s_i \le y_i$$

$$-\langle w, X_i\rangle - s_i \le -y_i$$

حال فرمول ابر صفحه‌ی بالا را در قالب ماتریس در می‌آوریم.

فرض کنید $A=\begin{bmatrix} X & -I_m \\ -X & -I_m \end{bmatrix}$ که $A\in\mathbb{R}^{2m\times(m+d)}$

فرض کنید $v\in\mathbb{R}^{m+d}$ یک برداری از متغیرهای $(w_1,\dots,w_d,s_1,\dots,s_m)$ باشد

تعریف می‌کنیم $b\in\mathbb{R}^{2m}$ که بردار $b=(y_1,\dots,y_m,-y_1,\dots,-y_m)^T$

در نهایت یک $c\in\mathbb{R}^{d+m}$ تعریف می‌کنیم که $c=(\underbrace{0,\dots,0}_{d},\underbrace{1,\dots,1}_{m})$ باشد.

این نشان می‌دهد که بهینه سازی مسئله‌ی ERM می‌تواند به صورت پیش بینی کننده های خطی نیز نشان داده شود:

$$\min\ c^T v \quad ; \quad Av \le b$$

۹.۶

فرض کنید $A=\{x_1,\dots,x_m\}\in\mathbb{R}^d$ توسط B_d، shatter می‌شود، پس برای هر $y=(y_1,\dots,y_d)\in\{-1,1\}^d$ وجود دارد $B_{\mu,r}\in B$ که برای هر i:

$$B_{\mu,r}(x_i)=y_i$$

بنابراین:

(☆) $$\left|\left((2\mu_i-1)^T\cdot(x_i;\|x_i\|^2)-\|\mu\|^2+r^2\right)\right|=y_i$$

برای هر $i\in[m]$: $\phi(x_i)=(x_i;\|x_i\|^2)$.

یک ابرصفحه $h\in L_{d+1}$ که وابسته به بردار ابرصفحه $w=(2\mu_i-1)$ و $b=\|\mu\|^2-r^2$ که در مدل (☆) را نتیجه می‌دهد پس:

$$h(x_i)=y_i$$

حال نتیجه می‌گیریم که اگر $A=\{x_1,\dots,x_m\}$ توسط B، shatter شود پس $\phi(A)=\{\phi(x_1),\dots,\phi(x_m)\}$ توسط L، shatter می‌شود. که نتیجه می‌دهد

$$VCdim(B_d)\leq VCdim(L_{d+1})$$

و

$$VCdim(L_{d+1})=d+2$$