课后练习2

1 问题一

我们要求的是：

$$\begin{aligned} &\underset{x,y,z}{min}\, x^2+y^2+z^2 \\ \text{s.t.} \quad &x+y+z=9 \\ &x+2y+3z=20 \end{aligned}$$

因此我们可以构造拉格朗日多项式：

$$L(x,y,z,\lambda_1,\lambda_2)=x^2+y^2+z^2+\lambda_1(x+y+z-9)+\lambda_2(x+2y+3z-20)$$

对于使$f(x,y,z)$最小的x^*,y^*,z^*，我们有：

$$\begin{aligned} \nabla_x L(x^*,y^*,z^*,\lambda_1,\lambda_2)=0 \\ \nabla_y L(x^*,y^*,z^*,\lambda_1,\lambda_2)=0 \\ \nabla_z L(x^*,y^*,z^*,\lambda_1,\lambda_2)=0 \\ \nabla_{\lambda_1} L(x^*,y^*,z^*,\lambda_1,\lambda_2)=0 \\ \nabla_{\lambda_2} L(x^*,y^*,z^*,\lambda_1,\lambda_2)=0 \end{aligned}$$

解得：

$$\begin{aligned} &x^*=2 \\ &y^*=3 \\ &z^*=4 \\ &\lambda_1=\lambda_2=-2 \end{aligned}$$

因此$\underset{x,y,z}{min}\, f(x,y,z)=29$

2 问题二

2.1

对于平面上的样本，它们都满足：

$$y_i(w^T x_i + b) \geq 1$$

取等号当且仅当x_i是支持向量. 现在已知$(6,5),(5,9),(2,4)$为支持向量，假设超平面为$w_1x_1 + w_2x_2 + b = 0$，则有：

$$\begin{aligned} y_{(6,5)}(6w_1 + 5w_2 + b) = 1 \\ y_{(5,9)}(5w_1 + 9w_2 + b) = 1 \\ y_{(2,4)}(2w_1 + 4w_2 + b) = 1 \end{aligned}$$

解得：

$$\begin{aligned} &w = (\frac{8}{17}, \frac{2}{17})^T \\ &b = -\frac{41}{17} \end{aligned}$$

2.2

对于$x_1 = (1,15)$和$x_2 = (4,2)$，有$w^T x_1 + b < 0$和$w^T x_2 + b < 0$. 因此两个新样本点都是负样本.

2.3

Proof 对于一个线性可分的数据分布，一定存在一个线性超平面$w^T x + b = 0$，使得正负样本能够完全位于超平面的两侧，其中支持向量是那些满足$y_i(w^T x_i + b) = 1$的向量.

我们不妨假设正或负样本中有至少一者不存在支持向量，即不存在使得$y_i(w^T x_i + b) = 1$的向量. 取正、负样本中$y_i(w^T x_i + b)$值最小的向量x_{j_1}和x_{j_2}并记其值为k_1和k_2，且$k_1 > k_2 \geq 1$.

将该平面向靠近x_{j_1}方向平移$\frac{k_1 - k_2}{2||w||}$的距离得到$w'^T x + b' = 0$，则此时有

$$\begin{aligned} &y_{j_1}(w'^T x_{j_1} + b') \\ =&y_{j_2}(w'^T x_{j_2} + b') \\ =&\frac{k_1 + k_2}{2} > 1 \end{aligned}$$

只需令

$$w'' = w'/(\frac{k_1 + k_2}{2})$$
$$b'' = b'/(\frac{k_1 + k_2}{2})$$

即可使x_{j_1}和x_{j_2}成为支持向量且此时有$||w''||^2 < ||w||^2$，这表明w''相比于w是一个更优解，说明假设不成立，则正负样本中均至少有一个支持向量.

3 问题三

3.1

由于已知$\alpha_i > 0, \ \forall i$，故这6个样本点均为支持向量，因此可以计算得到：

$$\begin{aligned} b &= \frac{1}{y_1} - \sum_{i=1}^{6} \alpha_i y_i \exp(-\frac{||x_i - x_1||^2}{2}) \\ &\approx -0.07389 \end{aligned}$$

3.2

$$f(x) = \sum_{i=1}^{6} \alpha_i y_i \exp(-\frac{||x - x_i||^2}{2}) + b$$

3.3

对于样本$x_4 = (2, 2)$：

$$f(x_4) = \sum_{i=1}^{6} \alpha_i y_i \exp(-\frac{||x_i - x_4||^2}{2}) + b < 0$$

故可以验证其是一个负样本

对于新的样本点$x_7 = (-2, -2)$：

$$f(x_7) = \sum_{i=1}^{6} \alpha_i y_i \exp(-\frac{||x_i - x_7||^2}{2}) + b > 0$$

故其是一个正样本

4 问题四

4.1

加入松弛变量前:

$$y_3(w^Tx_3+b)=1$$
$$y_4(w^Tx_4+b)=1$$

因此x_3和x_4是支持向量

加入松弛变量后，由于$\xi_1=\xi_2=\xi_4=0$，而

$$y_3({w'}^Tx_3+b')\geq 1-\xi_3=-0.5$$

因此这个问题转化为：

$$\begin{aligned}
&\min \frac{1}{2}||w||^2+C\sum_{i=1}^{4}\xi_i\\
\text{s.t.}\quad & w^Tx_3+b=-0.5\\
& w^Tx_4+b=-1\\
& w^Tx_1+b\geq 1\\
& w^Tx_2+b\geq 1
\end{aligned}$$

解得：

$$\begin{aligned}
&w=(\frac{1}{2},\frac{1}{2})^T\\
&b=-2
\end{aligned}$$

故新的分隔超平面为：

$$0.5x^{(1)}+0.5x^{(2)}-2=0$$

此时x_2, x_3和x_4为支持向量

4.2

硬间隔为：

$$\frac{|w^Tx_3+b|}{||w||}=\frac{\sqrt{2}}{4}$$

软间隔为：

$$\frac{|{w'}^Tx_2+b'|}{||w'||}=\sqrt{2}$$

间隔变大，模型的鲁棒性更高，对噪声的抗干扰能力更强，对新加入的样本也不容易误判

4.3

设四个数据点的Hinge Loss为$L_1, \ldots, L_4$，则：

$$
\begin{aligned}
L_1 &= \max\{0, 1 - y_1({w'}^T x_1 + b')\} = 0 \\
L_2 &= \max\{0, 1 - y_2({w'}^T x_2 + b')\} = 0 \\
L_3 &= \max\{0, 1 - y_3({w'}^T x_3 + b')\} = 1.5 \\
L_4 &= \max\{0, 1 - y_4({w'}^T x_4 + b')\} = 0
\end{aligned}
$$

只有分错的点和非常靠近超平面的点的Hinge Loss才不会为0，其余数据点均为0，不用考虑，即只有Loss不为0的点才会对模型产生影响，而这些点相对于整个样本空间是很少的，因此模型的解的稀疏性得到了保证