Solution Prob Set 4

Ex 8.1

I ran the code (modified from Jason Rebello's in Mathworks file exchange) with $\lambda = 0.001, 0.01, 0.1, 1, 10$.

I did not see big dependence of λ. Still if I followed the best λ from 5 fold cross validation, these are the numbers I got for error rates

	CV training	test	λ_{opt}
normalized	0.077	0.087	[illegible]
Log	0.056	0.057	10
binary	0.070	0.072	1

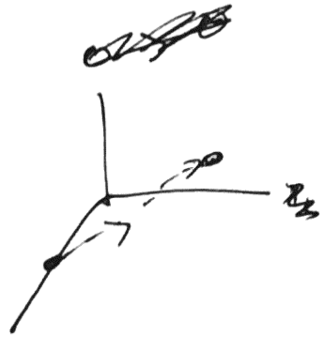
In different runs, the shallow min error λ ~~could~~ came out at a different value. The error rates were roughly the same.

Ex 14.1

a) $\phi(x_1) = \begin{pmatrix}1\\0\\0\end{pmatrix}$ $\quad \phi(x_2) = \begin{pmatrix}1\\2\\2\end{pmatrix}$

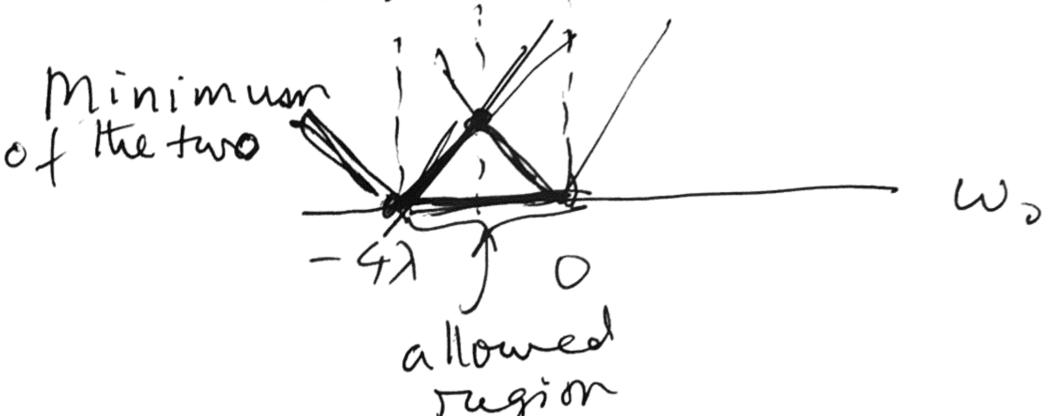
The vector separating them is $\begin{pmatrix}0\\1\\1\end{pmatrix}$

$$\omega = \lambda\begin{pmatrix}0\\1\\1\end{pmatrix}$$



$$\omega^T\phi(x_1) = 0$$
$$\omega^T\phi(x_2) = 4\lambda$$

b) Depending on w_0, Pt 1, is at $\frac{|w_0|}{\|w\|}$ and pt. 2 is at $\frac{|4\lambda + w_0|}{\|\omega\|}$



Best $\omega_0 = -2\lambda$

c) If we insist that $y(\omega^T\phi(x) + w_0) \geq 1$

Then $2\lambda \geq 1$ or $\lambda \geq +\frac{1}{2}$

Since margin $\frac{1}{\|w\|} = \frac{1}{\sqrt{2}\lambda}$, we have margin $= \sqrt{2}$

d) The two vectors are:

$$\phi(x_1) = (1, 0, 0)^T \qquad y_1 = -1$$

$$\phi(x_2) = (1, 2, 2)^T \qquad y = +1$$

$$\min \|w\|^2 \quad \text{s.t.} \quad \begin{aligned} &y_1 (w^T \phi(x_1) + w_0) \geq 1 \\ &y_2 (w^T \phi(x_2) + w_0) \geq 1 \end{aligned}$$

So minimize $w_1^2 + w_2^2 + w_3^2$

subject to

$$w_1 + w_0 \leq -1$$

$$w_1 + 2(w_2 + w_3) + w_0 \geq 1$$

$$\min_{w} \max_{\lambda \geq 0} \Big[w_1^2 + w_2^2 + w_3^2 - \lambda_1 (-w_1 - w_0 - 1) - \lambda_2 (w_1 + 2w_2 + 2w_3 + w_0 - 1) \Big] \qquad \lambda_1, \lambda_2 \geq 0$$

First w_0 variation gives $\lambda_1 = \lambda_2 = \lambda$

$$\max_{\lambda \geq} \min_{w} \Big[w_1^2 + w_2^2 + w_3^2 - 2\lambda (w_2 + w_3 - 1) \Big]$$

$\nabla_w \Rightarrow 0 \quad \Rightarrow \quad w_1 = 0$
$w_2 = \lambda$
$w_3 = \lambda$

With only two ~~points~~ points, one positive and the other negative, both are support vectors.

So.
$$y_1 w^T \phi(x_1) + w_0 = 1$$
$$y_2 w^T \phi(x_1) + w_0 = 1$$

$$\Rightarrow \quad (-1)\left[(0, \lambda, \lambda)\begin{pmatrix}1\\0\\0\end{pmatrix} + w_0\right] = 1 \Rightarrow \boxed{w_0 = -1}$$

and
$$(+1)\left[(0, \lambda, \lambda)\begin{pmatrix}1\\2\\2\end{pmatrix} + w_0\right] = 1 \Rightarrow 4\lambda + w_0 = 1$$

$$\boxed{\lambda = \frac{1}{2}}$$

$$w_0 = -1 \qquad \begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix} = \frac{1}{2}\begin{pmatrix}0\\1\\1\end{pmatrix}$$

e) Discriminant function

$$f(x) = w_0 + w^T \phi(x)$$
$$= -1 + \frac{x}{\sqrt{2}} + \frac{x^2}{2}$$