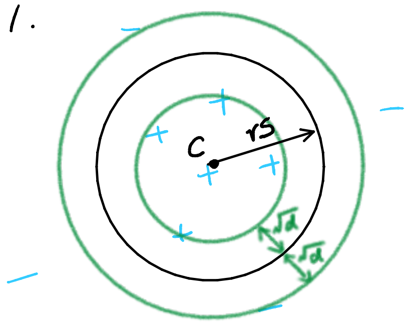
Problem Set3 Problem2

1.



The black circle is a spherical hypothesis. C is the center, $rs = r^2$ is the square of radius. Margin is $\sqrt{d}$.

So the optimization problem is:

$$\max_{c, rs, d} \quad d$$

$$\text{s.t.} \quad y^{(m)}\left(rs - \|x^{(m)} - c\|^2\right) \geq d \quad \text{for all } m.$$

$$rs \geq 0$$

(Since the object is maximizing d, I think it's ok not to include constraint $d \geq 0$.)

$$\Leftrightarrow \min_{c, rs, d} \ -d$$

$$\text{s.t.} \quad d - y^{(m)}\left(rs - \|x^{(m)} - c\|^2\right) \leq 0 \quad \text{for all } m.$$

$$-rs \leq 0$$

2. Using method of Lagrange multiplier:

$$L(x^{(m)}, y^{(m)}, c, rs, d)$$

$$= -d + \sum_m \lambda_m \left[d - y^{(m)}\left(rs - \|x^{(m)} - c\|^2\right)\right] + \xi(-rs)$$

Since the objective function is convex, the strict inequality constraint holds for at least one data point, the Slater's condition is satisfied.

Therefore, there's no gap between the dual problem and the primal problem.

To construct the dual problem:

Let $\frac{\partial L}{\partial d} = 0 \Leftrightarrow -1 + \sum_m \lambda_m = 0 \Leftrightarrow \boxed{\Sigma \lambda_m = 1}$

Let $\frac{\partial L}{\partial rs} = 0 \Leftrightarrow \sum_m \lambda_m(-y^{(m)}) - \xi = 0$

Let $\frac{\partial L}{\partial rs} = 0 \iff \sum_m \lambda_m (-y^{(m)}) - \xi = 0$

$$\iff \boxed{\sum_m \lambda_m y^{(m)} = -\xi}.$$

Before $\frac{\partial L}{\partial c}$, let's simplify L a little bit:

$$L = \cancel{-d} + \cancel{d} - \sum_m \lambda_m y^{(m)} (rs - \|x^{(m)} - c\|^2) + \sum_m \lambda_m y^{(m)} rs$$

$$= -\sum_m \cancel{\lambda_m y^{(m)}} rs + \sum_m \lambda_m y^{(m)} \|x^{(m)} - c\|^2 + \sum_m \cancel{\lambda_m y^{(m)}} rs$$

$$= \sum_m \lambda_m y^{(m)} (x^{(m)} - c)^T (x^{(m)} - c)$$

So, $\frac{\partial L}{\partial c} = -2 \sum_m \lambda_m y^{(m)} (x^{(m)} - c)$

$$\frac{\partial L}{\partial c} = 0 \iff \sum_m \lambda_m y^{(m)} (x^{(m)} - c) = 0$$

$$\iff \sum_m \lambda_m y^{(m)} x^{(m)} - \sum_m \lambda_m y^{(m)} c = 0 \quad (1)$$

$$\iff c = \frac{\sum_m \lambda_m y^{(m)} x^{(m)}}{\sum_m \lambda_m y^{(m)}} \quad (2)$$

Then we have

$$\inf_{c, rs, d} L = \sum_m \lambda_m y^{(m)} x^{(m)T} x^{(m)} - 2 \sum_m \lambda_m y^{(m)} x^{(m)T} c + \sum_m \lambda_m y^{(m)} c^T c$$

$$\overset{(1)}{=} \sum_m \lambda_m y^{(m)} x^{(m)T} x^{(m)} - \sum_m \lambda_m y^{(m)} x^{(m)T} c$$

$$\overset{(2)}{=} \sum_m \lambda_m y^{(m)} x^{(m)T} x^{(m)} + \frac{\sum_m \sum_n \lambda_m \lambda_n y^{(m)} y^{(n)} x^{(m)T} x^{(n)}}{\sum_m \lambda_m y^{(m)}}$$

So the dual problem is

$$\max_{\lambda, \xi} \sum_m \lambda_m y^{(m)} x^{(m)T} x^{(m)} + \frac{\sum_m \sum_n \lambda_m \lambda_n y^{(m)} y^{(n)} x^{(m)T} x^{(n)}}{\sum_m \lambda_m y^{(m)}}$$

$$\text{s.t.} \begin{cases} \lambda_m \geq 0 \\ \xi \geq 0 \\ \sum_m \lambda_m = 1 \\ \sum_m \lambda_m y^{(m)} = -\xi. \end{cases}$$