MidSem Exam Solutions

Ans 1 Let $x_1, x_2, \ldots, x_m$ be arbitrary points in $\mathcal{X}$.

Let K, K_1, K_2 denote the gram matrix with these points induced by k, k_1, k_2 respectively. Then,

$$K = \rho_1 K_1 + \rho_2 K_2 .$$

We need to show that K is psd

(a) K is symmetric because K_1, K_2 are.

(b) $\alpha^T K \alpha = \rho_1 \underbrace{\alpha^T K_1 \alpha}_{\substack{\geq 0 \\ (\because k_1 \text{ is kernel})}} + \rho_2 \underbrace{\alpha^T K_2 \alpha}_{\substack{\geq 0 \\ (\because k_2 \text{ is kernel})}} \geq 0 \quad \forall \alpha \in \mathbb{R}^m$

Hence k is a kernel.

Alternatively, you can answer (2) and prove (1)

Ans 2 Consider $\mathcal{H} \equiv \mathcal{H}_1 \underset{\text{direct sum}}{\oplus} \mathcal{H}_2$

and $\varphi(x) \equiv \left(\sqrt{\rho_1}\varphi_1(x), \sqrt{\rho_2}\varphi_2(x)\right)$ (ordered pairs)

It is easy to see that this works!

because,

$$k(x,z) = \langle \varphi(x), \varphi(z) \rangle_{\mathcal{H}} = \langle \sqrt{\rho_1}\varphi_1(x), \sqrt{\rho_1}\varphi_1(z) \rangle_{\mathcal{H}_1} + \langle \sqrt{\rho_2}\varphi_2(x), \sqrt{\rho_2}\varphi_2(z) \rangle_{\mathcal{H}_2}$$

($\because$ direct sum)

$$= \rho_1 k_1(x,z) + \rho_2 k_2(x,z)$$

Archer
PAGE NO.
DATE: / /

(3a) $$\min_{\omega \in \mathcal{H}} \frac{1}{2}\|\omega\|_{\mathcal{H}}^2 + \frac{C}{m}\sum_{i=1}^{m} \xi_i \qquad \text{(I)}$$

s.t. $\langle \omega, \varphi(x_i) \rangle_{\mathcal{H}} \geq 1 - \xi_i, \ \xi_i \geq 0 \ \forall\, i = 1, \ldots, m$

When φ & C are chosen appropriately such that approximation error is near zero, then we use $\langle \omega, \varphi(x) \rangle_{\mathcal{H}}$ as a proxy for density at x.

(3b) $$\varphi(x) = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2}\, x_1 x_2 \end{bmatrix}, \text{ where } x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(3c) (I) simplifies to:

$$\min_{\omega \in \mathbb{R}^3} \frac{1}{2}(\omega_1^2 + \omega_2^2 + \omega_3^2) + \frac{C}{7}\sum_{i=1}^{7} \xi_i$$

s.t.
$$\begin{aligned}
0 &\geq 1 - \xi_1, \ \xi_1 \geq 0 \\
\omega_1 &\geq 1 - \xi_2, \ \xi_2 \geq 0 \\
\omega_2 &\geq 1 - \xi_3, \ \xi_3 \geq 0 \\
\omega_1 &\geq 1 - \xi_4, \ \xi_4 \geq 0 \\
\omega_2 &\geq 1 - \xi_5, \ \xi_5 \geq 0 \\
100\omega_1 &\geq 1 - \xi_6, \ \xi_6 \geq 0 \\
100\omega_2 &\geq 1 - \xi_7, \ \xi_7 \geq 0
\end{aligned}$$

It is clear that $\omega_3^* = 0$ at optimality, and further by eliminating ξ_i (as in lectures) we obtain:

$$\text{I} = \min_{\omega_1 \in \mathbb{R}} \frac{1}{2}\omega_1^2 + \frac{C}{7}\Big(2\max(0, 1-\omega_1) + \max(0, 1-100\omega_1)\Big)$$

$$+ \underbrace{\min_{\omega_2 \in \mathbb{R}} \frac{1}{2}\omega_2^2 + \frac{C}{7}\Big(2\max(0, 1-\omega_2) + \max(0, 1-100\omega_2)\Big)}_{\text{II}}$$

$$+ \frac{C}{7}.$$

Again, it is clear that $\omega_1^* = \omega_2^*$. Hence it is enough to focus on II.

$$\text{II} = \min\Big(\underbrace{\min_{\omega_2 \le \frac{1}{100}} g(\omega_2)}_{\text{IIa}}, \underbrace{\min_{\frac{1}{100} < \omega_2 \le 1} g(\omega_2)}_{\text{IIb}}, \underbrace{\min_{\omega_2 > 1} g(\omega_2)}_{\text{IIc}}\Big)$$

$$\text{IIa} = \min_{\omega_2 \le \frac{1}{100}} \frac{1}{2}\omega_2^2 + \frac{2C}{7}(1-\omega_2) + \frac{C}{7}(1-100\omega_2)$$

In this case, if $\frac{1}{100} \ge \frac{102C}{7}$, then

$\omega_2^* = \frac{102C}{7}$, and $g(\omega_2^*) = \left(6C - \frac{102^2 C^2}{7}\right)\frac{1}{14}$

if $\frac{1}{100} \le \frac{102C}{7}$, then $\omega_2^* = \frac{1}{100}$, $g(\omega_2^*) =$

$$\frac{1}{14}\left(\frac{7}{100^2} - \frac{204C}{100} + 6C\right)$$

$$\text{IIb} = \min_{\frac{1}{100} < \omega_2 \leq 1} \frac{1}{2}\omega_2^2 + \frac{2c}{7}(1-\omega_2)$$

here, if $\frac{1}{100} < \frac{4c}{7} < 1$, then $\omega_2^* = \frac{2c}{7}$

and $g(\omega_2^*) = (4c - 4c^2/7)\frac{1}{14}$

else if $\frac{1}{100} \geq \frac{4c}{7}$, then $\omega_2^* = \frac{1}{100}$,

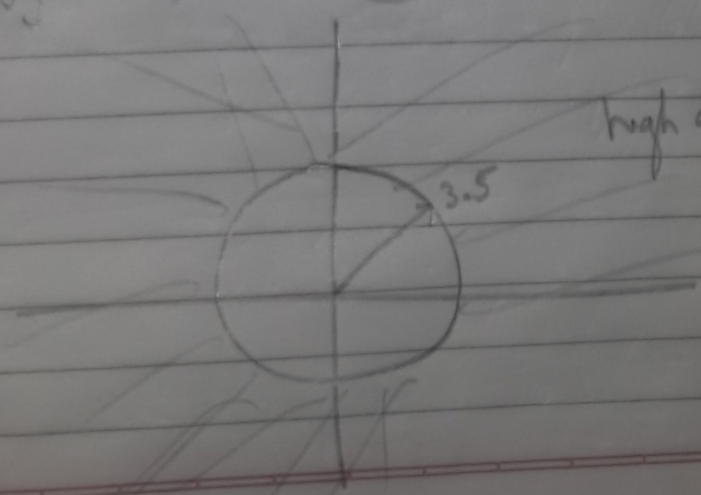
$$g(\omega_2^*) = \frac{1}{14}\left(\frac{7}{100^2} - \frac{4c}{100} + 4c\right)$$

else if $\frac{4c}{7} \geq 1$, then $\omega_2^* = 1$,

$$g(\omega_2^*) = 1/2.$$

$$\text{IIc} = \min_{\omega_2 \geq 1} \frac{1}{2}\omega_2^2 \quad \text{Here } \omega_2^* = 1,\ g(\omega_2^*) = 1/2$$

3d) If $c = 1$, then IIb is the minimum, hence $\omega_2^* = \frac{2}{7}$



high density

(3e) From the training data one might guess that it is very likely that high density regions are those within a circle/ellipse, than those outside, which is given by the kernel. Hence it seems to be not well suited.

(3f) With k, we will be looking at

$$\omega_1 x_1^2 + \omega_2 x_2^2 + \sqrt{2} \geq 1$$

$$\Leftrightarrow (-\omega_1) x_1^2 + (-\omega_2) x_2^2 \leq -0.44\ldots$$

$\downarrow$

This looks like regions inside a circle/ellipse.

Hence this is a much well-suited kernel.