# KMML Homework 3

Mahdi Kallel

TOTAL POINTS

## 4 / 4

QUESTION 1

### 1 Question 1 1 / 1

✓ **- 0 pts** **Correct**

**- 1 pts** wrong or missing answer

**- 0.25 pts** missing absolute value

**- 0.5 pts** missing final step

QUESTION 2

### 2 Question 2 3 / 3

✓ **- 0 pts** **Correct**

**- 0.5 pts** Representer theorem not or wrongly justified

**- 0.5 pts** K_X and K_Y are not necessarily invertible.

**- 1.5 pts** Computation not finished: what is the solution to this optimisation problem?

**- 2 pts** Wrong formula

**- 1 pts** Computation not finished: what is the solution to this optimisation problem?

**- 3 pts** Problem not solved

**- 0.5 pts** Computation not finished: what is the solution to this optimisation problem?

Kernel Methods DM#3 :

Exercice 1:

$The\ RKHS\ of\ the\ linear\ kernel\ is\ \{f_w(x) =< w, x > \ \forall w \ \in \mathbb{R}\ \}\ and\ thus\ it\ is\ the\ scalar\ multiplication.$

$By\ this\ definition\ \exists f, g \ \in \mathbb{R}\ s.t\ :$

$$cov_n(f(X), g(Y)) = \frac{1}{n}\sum_i f.x_i * g.y_i - \frac{1}{n^2}\sum_i f.x_i * \sum_i g.y_i.$$

$$= \frac{fg}{n}\left(X^{\ T}Y - X^T\, O\, Y\right) = \frac{fg}{n} X^{\ T}(I_n - O)\, Y\ ,\ where\ O = \frac{\left(\mathbb{1}\mathbb{1}^T\right)}{n}$$

$From\ the\ unit\ ball\ constraint\ we\ have\ that\ |f|, |g| \leq 1.\ And\ thus\ we\ deduce\ that\ :$

$$C_N^K(X, Y) = max_{|f, g| \leq 1} \frac{fg}{n} X^{\ T}(I_n - O)\, Y = \frac{|X^{\ T}(I_n - O)\, Y|}{n}$$

1 Question 1 **1 / 1**

✓ **- 0 pts** **Correct**

**- 1 pts** wrong or missing answer

**- 0.25 pts** missing absolute value

**- 0.5 pts** missing final step

Exercice 2 :

*We start by showing that we can restrict our* $f, g$ *solutions to* $C_n^k$ *to the form* $f = \sum \alpha_i K_{x_i}$, $g = \sum \beta_i K_{y_i}$ .

*Let's* $\mathcal{H}_x = \left\{ f \text{ s.t } f = \sum \alpha_i K_{x_i}\ (\alpha_1 .. \alpha_n) \in \mathbb{R}^n \right\}$ $\mathcal{H}_x$ *is a finite dimensional vector space.*

*Therefore* $\forall f \in \mathcal{H},\ f = f_x + f_\perp$.

$\forall_i\ f(x_i) = < f_x, K_{xi} >$ *and by orthogonality* $||f_x||^2 = ||f||^2 - ||f_\perp||^2$ *therefore* $||f_x|| \le ||f||$.

*Thus for any solution* $f^\star$ *of* $C_n^k$ *we can find a projection* $f_x^\star$ *that can be written as* $\sum \alpha_i K_{x_i}$ *and is also a solution of* $C_n^k$.

$$cov_n(f(X), g(Y)) = \frac{1}{n}\sum_i f(u_i) * g(y_i) - \frac{1}{n^2}\sum_i f(x_i) * \sum_i g(y_i)$$

$$\frac{1}{n}\sum_i [K_x F]_i [K_y G]_i - \frac{1}{n}\sum_i [K_x F]_i \frac{1}{n}\sum_i [K_y G]_i$$

$$= \frac{1}{n}(K_x F)^T (K_y G) - \frac{1}{n}(K_x F)^T O\, K_y G = \frac{1}{n}(K_x F)^T (I_n - O)(K_y G)$$

*Since the representer theorem applies our norm constraints translate to* :

$F^T K_x F,\ G^T K_y G \le 1.$

$$And\ thus\ C_n^k(X,Y) \;=\; max_{F,G}\ \frac{1}{n}\ (K_xF)^T\ (I_n - O)\ (K_yG)$$

$$s.t:\ \ F^TK_xF,\ G^TK_yG\ \le 1$$

$K_x, K_y$ *are positive semi definite and thus have a root* $\sqrt{K_x},\ \sqrt{K_y}$

$$C_n^k(X,Y) \;=\; max_{F,G}\ \frac{1}{n}\ \left(F^TK_x\right)\ (I_n - O)\ (K_yG)\ \ st...$$

$$=\; max_{F,G}\ \frac{1}{n}\ F^T\sqrt{K_x}\ \sqrt{K_x}\ (I_n - O)\ \sqrt{K_y}\sqrt{K_y}G$$

$$s.t\ \ ||F^T\sqrt{K_x}||^2,\ ||G^T\sqrt{K_y}||^2 \le 1\ \ \leftrightarrow\ \ ||F^T\sqrt{K_x}||\ \ ,\ ||G^T\sqrt{K_y}|| \le 1$$

*We now want to show that this is equivalent to the following problem* :

$$C_n^k(X,Y) \;=\; max_{\widetilde{F},\widetilde{G}}\ \frac{1}{n}\ \widetilde{F}^T\ \sqrt{K_x}\ (I_n - O)\ \sqrt{K_y}\ \widetilde{G}$$

$$s.t\ \ ||\widetilde{F}||,\ ||\widetilde{G}||\ \le 1$$

$\rightarrow$ *If* $F, G$ *are solutions of the original problem, then we can define* $\widetilde{F} \;=\; \sqrt{K_xF}\ ,\ \widetilde{G} = \sqrt{K_yG}$
*and we get* $||\widetilde{F}||,\ ||\widetilde{G}||\ \le 1$ *and* $\frac{1}{n}\ \widetilde{F}^T\ \sqrt{K_x}\ (I_n - O)\ \sqrt{K_y}\ \widetilde{G} \;=\; \frac{1}{n}\ \left(F^TK_x\right)\ (I_n - O)\ (K_yG).$

$\leftarrow$ *If* $\widetilde{F}, \widetilde{G}$ *are solutions of the second problem, since* $K_x, K_y$ *are p.d it's diagonalizable in an orthogonal*

$$\Longrightarrow E \;=\; Im(K_x) \bigoplus Ker(K_x) \;=\; Im(K_y) \bigoplus Ker(K_y)$$

$\exists F\ , F_k\ st\ \ \widetilde{F} \;=\; \sqrt{K_x}F\ \ +\ F_k$ *and* $\sqrt{K_x}F_k = 0.$ *(and the same for* $\widetilde{G}$*)*

$$\frac{1}{n}\ \widetilde{F}^T\ \sqrt{K_x}\ (I_n - O)\ \sqrt{K_y}\ \widetilde{G} \;=\; \frac{1}{n}\ F^T\sqrt{K_x}\sqrt{K_x}\ (I_n - O)\ \sqrt{K_y}\ \sqrt{K_y}G$$

$$= \frac{1}{n} \; F^T K_x \, (I_n - O) \, K_y G.$$

*By orthogonality* : $||F||^2 \; = \; ||\widetilde{F}||^2 - ||F_k||^2 \quad \leq 1 \; - ||F_k||^2 \leq 1$

*Thus we've shown that our optimization problem can be rewritten as* :

$$C_n^k(X, Y) \;\; = \;\; max_{\widetilde{F},\widetilde{G}} \frac{1}{n} \widetilde{F}^T \; \sqrt{K_x} \, (I_n - O) \, \sqrt{K_y} \, \widetilde{G}$$

$$s.t \;\; ||\widetilde{F}||, \; ||\widetilde{G}|| \;\; \leq 1$$

*For a fixed F , we call* $A_G \;\; = \;\; \sqrt{K_x} \, (I_n - O) \sqrt{K_y} \, \widetilde{G}$

*The problem is then* : $max_{||\widetilde{F}|| \leq 1} \; \widetilde{F}^T A_G \;\; = max_{\widetilde{F}} \; \frac{\widetilde{F}^T A_G}{||\widetilde{F}||}$

*which is solved for* $\widetilde{F} \;\; = \;\; \frac{A_G}{||A_G||}$ *(using cauchy schwarz).*

*And thus the problem can be rewritten as* :

$$max_{\;||\widetilde{G}|| \leq 1} \; \frac{A_G^T A_G}{||A_G||} \; = max_{\;||\widetilde{G}|| \leq 1} \; || \, A_G|| \;\; = \;\; max_{\;||\widetilde{G}|| \leq 1} || \, \sqrt{K_x} \, (I_n - O) \sqrt{K_y} \, \widetilde{G}||$$

*One can show that* : $max_{\;||\widetilde{G}|| \leq 1} || \, \sqrt{K_x} \, (I_n - O) \sqrt{K_y} \, \widetilde{G}|| \;\; = \;\; max_{\;||\widetilde{G}|| =1} || \, \sqrt{K_x} \, (I_n - O) \sqrt{K_y} \, \widetilde{G}||$

$$\left( \textit{It suffices to multiply any solution } G^* \textit{ by } \frac{1}{||G^*||} \textit{ to get a better solution} \right)$$

*Thus we recognize the spectral norm problem and thus* :

$C_n^k(X, Y) \;\; = \;\; \frac{\lambda_{max}}{n}$ *where* $\lambda_{max}$ *is the maximum eigenvalue of* $\sqrt{K_x} \, (I_n - O) \sqrt{K_y}$ .

## 2 Question 2 **3 / 3**

✓ **- 0 pts** **Correct**

**- 0.5 pts** Representer theorem not or wrongly justified

**- 0.5 pts** K_X and K_Y are not necessarily invertible.

**- 1.5 pts** Computation not finished: what is the solution to this optimisation problem?

**- 2 pts** Wrong formula

**- 1 pts** Computation not finished: what is the solution to this optimisation problem?

**- 3 pts** Problem not solved

**- 0.5 pts** Computation not finished: what is the solution to this optimisation problem?