1.a

$k(x,y) = \frac{1}{1-xy}$ را داریم. حال بسط تیلور این تابع را می‌نویسیم

$x,y \in (-1,1) \longrightarrow xy \in (-1,1)$

$\longrightarrow k(x,y) = \sum_{k=0}^{\infty} (xy)^k$

حال می‌دانیم که $k(x,y) = xy$ یک کرنل pds است و اگر به توان k برسد نیز
کرنل چندجمله‌ای می‌شود که در کلاس گفته شد و می‌دانیم pds است. در نهایت نیز می‌دانیم جمع یکسری کرنل pds همچنان pds
است $\longleftarrow$ $k(x,y) = \frac{1}{1-xy}$ کرنل pds است

1.b

می‌دانیم که اگر یک کرنل pds باشد باید ماتریس کرنل $[k(x_i, x_j)]_{n\times n}$ psd باشد
بنابراین اگر بتوانیم یک دیتاست n تایی (به ازای هر n) مثال بزنیم که این ماتریس psd نشود آنگاه کرنل pds نخواهد بود

مثال نقض $\{x_1 = 2 ,\ x_2 = 3\}$ $\quad n = 2$

$$k_x = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) \\ k(x_2, x_1) & k(x_2, x_2) \end{bmatrix} = \begin{bmatrix} \lg 5 & \lg 7 \\ \lg 7 & \lg 10 \end{bmatrix} \longrightarrow \det \simeq -0.015$$

می‌دانیم که اگر det منفی شود یعنی حداقل یکی از eig val ها منفی است و ماتریس psd نیست
بنابراین مثال نقضی ارائه کردیم که k_x برای آن psd نبود $\longleftarrow$ $k(x,y) = \lg(1+xy)$ کرنل pds نیست

$\boxed{1.c}$

مشابه سوال قبل مثال نقض ارائه می دهیم که k_X به ازای آن psd نباشد

مثال نقض $\left\{x_1 = \frac{3\pi}{4} \ , \ x_2 = \frac{-3\pi}{4}\right\}$ $\quad n=2$

$$k_X = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) \\ k(x_2, x_1) & k(x_2, x_2) \end{bmatrix} = \begin{bmatrix} \cos\left(\frac{3\pi}{2}\right) & \cos(0) \\ \cos(0) & \cos\left(\frac{-3\pi}{2}\right) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow{\det} -1$$

بنابراین مثال نقض ارائه کردیم که k_X برای آن psd نبود $\longrightarrow$ کرنل $k(x,y) = \cos(x+y)$ pds نیست

$\boxed{1.d}$

در حالتی که $x, y \in \mathbb{R}_+$ باشند داریم :

$$\int_0^\infty \mathbb{1}_{t\in[0,x]} \cdot \mathbb{1}_{t\in[0,y]} \, dt = \int_0^{\min(x,y)} 1 \, dt = \min(x,y)$$

بنابراین معلوم شد که اگر $\Phi(x) = \mathbb{1}_{t\in[0,x]}$ قرار دهیم با توجه به تعریف ضرب داخلی برای توابع اندازه پذیر ،

کرنل $k(x,y) = \min(x,y)$ یک کرنل pds است زیرا برابر $\langle \Phi(x) , \Phi(y) \rangle$ است .

2.1

طبق الگوریتم اگر $\| \Phi(x) - c_+ \| \leqslant \| \Phi(x) - c_- \|$ باشد $h(x) = 1$ و در غیر این صورت $h(x) = -1$

$$\rightarrow \| \Phi(x) - c_+ \|^2 \leqslant \| \Phi(x) - c_- \|^2$$

$$\rightarrow \cancel{\| \Phi(x) \|^2} - 2 \langle \Phi(x), c_+ \rangle + \| c_+ \|^2 - \cancel{\| \Phi(x) \|^2} + 2 \langle \Phi(x), c_- \rangle - \| c_- \|^2 \leqslant 0$$

$$\rightarrow 2 \langle \Phi(x), \underbrace{c_+ - c_-}_{w} \rangle + \underbrace{\left(\| c_- \|^2 - \| c_+ \|^2 \right)}_{2b} \geqslant 0$$

بنابراین اگر w و b را به صورت مشخص شده تعیین کنیم

واضح است که می توان نوشت $h(x) = \mathrm{sign}\left(\langle \Phi(x), w \rangle + b \right)$

2.2

مقادیر c_+ و c_- را در فرمول $h(x)$ جایگذاری می کنیم

$$h(x) = \mathrm{sign}\left(\underbrace{\langle \Phi(x), c_+ \rangle}_{\frac{1}{m_+} \sum_{i: y_i = +1} k(x_i, x)} - \langle \Phi(x), c_- \rangle + \frac{1}{2} \underbrace{\| c_- \|^2}_{\frac{1}{m_-^2} \sum_{\substack{i: y_i = -1 \\ j: y_j = -1}} k(x_i, x_j)} - \frac{1}{2} \| c_+ \|^2 \right)$$

بنابراین نشان دادیم که هر 4 ترم مورد نیاز برای $h(x)$ تنها به $k(\cdot, \cdot)$ نیاز دارد پس محاسبه $h(x)$ تماماً بر اساس تابع کرنل است

$\boxed{3.1}$

می دانیم که هر تابع مانند f که عضو RKHS باشد را می توان به صورت ترکیب خطی پایه های $\Phi(x_i)$ نوشت

$$f(x) \triangleq \sum_{i\in I} \underbrace{a_i \underbrace{\Phi(x_i)(x)}_{x_i x}}_{\beta_i} = \sum_{i\in I} \beta_i x = x \underbrace{\sum \beta_i}_{\beta_f}$$

$$\langle f, f \rangle_{\mathbb{H}} = \sum_{i,j\in I,J} a_i a_j \underbrace{k(x_i, x_j)}_{x_i x_j} = \Big(\underbrace{\sum_{i\in I} \overbrace{a_i x_i}^{\beta_i}}_{\beta_f}\Big)^2 \longrightarrow \|f\|_{\mathbb{H}} = |\beta_f|$$

$$\longrightarrow C_n^k(X, Y) = \max_{f,g\in B_k} \operatorname{cov}_n(\beta_f X, \beta_g Y)$$

$$= \max_{\substack{\beta_f \leq 1, \\ \beta_g \leq 1}} \beta_f \beta_g \operatorname{cov}_n(X, Y)$$

$$= \left| \operatorname{cov}_n(X, Y) \right|$$

$\boxed{3.2}$ $$c_n^k(X,Y) = \max_{f,g\in B_k} \frac{1}{n}\sum_{i=1}^{n} f(x_i)g(y_i) - \left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right)\left(\frac{1}{n}\sum_{i=1}^{n} g(y_i)\right)$$

$$= \min_{f,g}\ \max_{\lambda_1,\lambda_2}\ \frac{-1}{n}\sum_{i=1}^{n} f(x_i)g(y_i) + \left(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\right)\left(\frac{1}{n}\sum_{i=1}^{n} g(y_i)\right)$$

$$+ \lambda_1\left(\|f\|_{\mathbb{H}} - 1\right) + \lambda_2\left(\|g\|_{\mathbb{H}} - 1\right)$$

فرض می‌کنیم
strong duality
برقرار است

$$= \max_{\lambda_1,\lambda_2}\ \underbrace{\min_{f,g}\ \text{عبارت بالا}}_{\mathrm{I}}$$

I به فرم representer theorem است . بنابر این داریم :

$$f^* = \sum_{j=1}^{n} \alpha_j k(x_j,\cdot)$$

$$g^* = \sum_{j=1}^{n} \beta_j k(y_j,\cdot)$$

f^* , g^* را جاگذاری می‌کنیم
و λ_1, λ_2 را دوباره به فرم قیدها
می‌نویسیم

$$\longrightarrow \min_{\alpha,\beta}\ -\frac{1}{n}\sum_{i=1}^{n}\left[\left(\sum_{j=1}^{n}\alpha_j k(x_j,x_i)\right)\left(\sum_{j=1}^{n}\beta_j k(y_j,y_i)\right)\right]$$

$$+\frac{1}{n^2}\left(\sum_{i,j=1}^{n}\alpha_j k(x_j,x_i)\right)\left(\sum_{i,j=1}^{n}\beta_j k(y_j,y_i)\right)$$

$$\text{s.t. } \|f\|_{\mathbb{H}} = \sum_{i,j}\alpha_i\alpha_j k(x_i,x_j) \leqslant 1$$

$$\|g\|_{\mathbb{H}} = \sum_{i,j}\beta_i\beta_j k(y_i,y_j) \leqslant 1$$

فرم ماتریسی

centering matrix $= \mathbb{H}$

$$= \max_{\alpha,\beta}\ \frac{1}{n}\alpha^T K_x\left(I - \frac{1}{n}1_n 1_n^T\right)K_y\,\beta$$

$$\text{s.t. } \alpha^T K_x\alpha \leqslant 1\ ,\ \beta^T K_y\beta \leqslant 1$$

$\boxed{3.2}$ ادامه

تعریف می‌کنیم $\tilde{\alpha} = k_x^{\frac{1}{2}} \alpha$, $\tilde{\beta} = k_y^{\frac{1}{2}} \beta$ و بنابراین داریم:

$$= \max_{\tilde{\alpha}, \tilde{\beta}} \tilde{\alpha}^T A \tilde{\beta} \qquad A = k_x^{\frac{1}{2}} H k_y^{\frac{1}{2}}$$

$$\text{s.t. } \|\tilde{\alpha}\| \leqslant 1 \ , \ \|\tilde{\beta}\| \leqslant 1$$

که می‌دانیم جواب maximization فوق $\lambda_{max}(A)$ است یعنی بیشترین مقدار ویژه ماتریس A

$\boxed{4.1}$

اگر نشان دهیم ماتریس کرنل به ازای هر n همواره symmetric و psd است آنگاه کرنل pds خواهد بود

از آنجا که $\mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B)\mathbb{P}(A)$ و $\mathbb{P}(A\cap B) = \mathbb{P}(B\cap A)$ واضح است که ماتریس کرنل symmetric است.

برای psd بودن k_x باید داشته باشیم:

$$\forall n \quad k_x = \left[k(x_i, x_j)\right]_{n\times n}$$

$$\forall \alpha\in\mathbb{R}^n \quad \alpha^T k_x \alpha \geqslant 0$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j \underbrace{k(x_i, x_j)}_{\mathbb{P}(x_i\cap x_j) - \mathbb{P}(x_i)\mathbb{P}(x_j)} \geqslant 0$$

$$\rightarrow \underbrace{\underbrace{\sum_{i,j}\alpha_i\alpha_j \underbrace{\mathbb{P}(x_i\cap x_j)}_{\mathbb{E}[1_{x_i} 1_{x_j}]}}_{\mathbb{E}\left[\underbrace{\sum_{i,j}\alpha_i\alpha_j 1_{x_i}1_{x_j}}_{(\sum_i \alpha_i 1_{x_i})^2}\right]}} - \underbrace{\underbrace{\sum_{i,j}\alpha_i\alpha_j \mathbb{P}(x_i)\mathbb{P}(x_j)}_{\left(\sum_i \alpha_i \underbrace{\mathbb{P}(x_i)}_{\mathbb{E}[1_{x_i}]}\right)^2}}_{\mathbb{E}\left[\sum_i \alpha_i 1_{x_i}\right]^2}$$

اگر $Z = \sum_{i=1}^{n}\alpha_i 1_{x_i}$ آنگاه عبارت فوق برابر $Var(Z)$ است.

و می دانیم که Var همواره بزرگتر مساوی 0 است $\longrightarrow$ k_x is psd

بنابراین به ازای هر n ماتریس کرنل symmetric و psd است $\longrightarrow$ کرنل pds است

$\boxed{4.2}$

در این سوال Φ را به صورت explicit معرفی می‌کنیم تا اثبات شود کرنل pDs است.

$p(s)$ دارای 2^m عضو است که آنها را $s'_1 \dots s'_{2^m}$ می‌نامیم.

حال Φ را به صورت زیر معرفی می‌کنیم:

$$\Phi(A) = \begin{bmatrix} \mathbb{1}_{s'_1 \subseteq A} \\ \vdots \\ \mathbb{1}_{s'_{2^m} \subseteq A} \end{bmatrix}_{2^m \times 1}$$

$$\rightarrow \Phi(A) \cdot \Phi(B) = \sum_{i=1}^{2^m} \mathbb{1}_{s'_i \subseteq A} \mathbb{1}_{s'_i \subseteq B} = \sum_{c \subseteq A \cap B} 1$$

$$= \sum_{i=\emptyset}^{|A \cap B|} \underbrace{\binom{|A \cap B|}{i}}_{\substack{\text{تعداد زیر مجموعه های } i \text{ تایی} \\ \text{از } A\cap B}} = 2^{|A \cap B|}$$

بنابراین Φ ای ارائه دادیم که $k(A,B) = \langle \Phi(A), \Phi(B) \rangle$

پس کرنل pDs است

4.3

برای حالتی که $x, x' \in \mathbb{R}$ باشند داریم :

$$\int_0^\infty \mathbb{1}_{t \in [0, |x|]} \ \mathbb{1}_{t \in [0, |x'|]} \ dt$$

$$= \int_0^{\min(|x|, |x'|)} 1 \ dt = \min(|x|, |x'|)$$

بنابر این معلوم شد که اگر $\Phi(x) = \mathbb{1}_{t \in [0, |x|]}$ قرار دهیم با توجه به تعریف ضرب داخلی برای توابع اندازه پذیر ،

کرنل $k(x, x') = \min(|x|, |x'|)$ یک کرنل PDS است زیرا برابر $\langle \Phi(x), \Phi(x') \rangle$ است .

همچنین واضح است که در تمامی فرمول های بالا به جای $|x|$ می توان $|x|^\alpha$ گذاشت پس به همین صورت $k_\alpha(x, x') = \min(|x|^\alpha, |x'|^\alpha)$

نیز یک کرنل PDS است .

حال می دانیم که جمع یکسری کرنل PDS همچنان PDS است $\longleftarrow$ $K_\alpha(x, x') = \sum_{k=1}^{N} \min(|x|^\alpha, |x'|^\alpha)$ $\quad x, x' \in \mathbb{R}^N$

نیز یک کرنل PDS است

در برخی سوال ها با علی پاپایک همفکری داشتم