Stats 231 / CS229T Homework 3 Solutions

Question 1: Let $\mathsf{k} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a valid kernel function. Define

$$\mathsf{k}_{\text{norm}}(x, z) := \frac{\mathsf{k}(x, z)}{\sqrt{\mathsf{k}(x, x)}\sqrt{\mathsf{k}(z, z)}}.$$

Is k_{norm} a valid kernel? Justify your answer.

Answer: Yes, it is. Let $\mathsf{k}(x, z) = \langle \phi(x), \phi(z) \rangle$ for some mapping $\phi : \mathcal{X} \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space. Then

$$\mathsf{k}_{\text{norm}}(x, z) = \langle \phi(x)/ \left\|\phi(x)\right\|_2, \phi(z)/ \left\|\phi(z)\right\|_2 \rangle$$

so that it is still a valid inner product, where the feature mapping is now $x \mapsto \phi(x)/ \left\|\phi(x)\right\|_2$ for $\left\|\phi(x)\right\|_2^2 = \langle \phi(x), \phi(x) \rangle$. □

Question 2: Consider the class of functions

$$\mathcal{H} := \left\{ f : f(0) = 0, f' \in L^2([0, 1]) \right\},$$

that is, functions $f : [0, 1] \to \mathbb{R}$ with $f(0) = 0$ that are almost everywhere differentiable, where $\int_0^1 (f'(t))^2 dt < \infty$. On this space of functions, we define the inner product by

$$\langle f, g \rangle = \int_0^1 f'(t) g'(t) dt.$$

Show that $\mathsf{k}(x, z) = \min\{x, z\}$ is the reproducing kernel for $\mathcal{H}$, so that it is (i) positive semidefinite and (ii) a valid kernel.

Answer: If we show that $\mathsf{k}(x, z) = \min\{x, z\}$ is indeed the reproducing kernel for $\mathcal{H}$, then that suffices to demonstrate that it is a positive definite function. We have for $g(z) = \mathsf{k}(x, z)$ that (almost everywhere) $g'(z) = \mathbf{1}\{x \le z\}$, so that

$$\langle f, \mathsf{k}(z, \cdot) \rangle = \int_0^1 f'(t) \mathbf{1}\{t \le z\} dt = \int_0^z f'(t) dt = f(z) - f(0) = f(z).$$

Thus k is evidently a reproducing kernel, so it must be a positive definite function.

(Another way to see that, we have $\min\{x, z\} = \mathsf{k}(x, z) = \int_0^1 \mathbf{1}\{t \le x\} \mathbf{1}\{t \le z\} dt$, so that $\min\{x, z\}$ is evidently an inner product.) □

Question 3: Consider the Sobolev space $\mathcal{F}_k$, which is defined as the set of functions that are $(k-1)$-times differentiable and have kth derivative almost everywhere on $[0, 1]$, where the kth derivative is square-integrable. That is, we define

$$\mathcal{F}_k := \left\{ f : [0, 1] \mid f^{(k)} \in L^2([0, 1]) \right\},$$

where $f^{(k)}$ denotes the kth derivative of f. We define the inner product on $\mathcal{F}_k$ by

$$\langle f, g \rangle = \sum_{i=0}^{k-1} f^{(i)}(0) g^{(i)}(0) + \int_0^1 f^{(k)}(t) g^{(k)}(t) dt.$$

(a) Find the representer of evaluation for this Hilbert space, that is, find a function $r_x : [0,1] \to \mathbb{R}$ (defined for each $x \in [0,1]$) such that $r_x \in \mathcal{F}_k$ and

$$\langle r_x, f \rangle = f(x)$$

for all $x \in [0,1]$.

(b) What is the reproducing kernel $\mathsf{k}(x,z)$ associated with this space? (Recall that $\mathsf{k}(x,z) = \langle r_x, r_z \rangle$ for an RKHS.)

(c) Show that $\mathcal{F}_k$ is a Hilbert space, meaning that $\|f\|^2 = \langle f, f \rangle$ defines a norm and that $\mathcal{F}_k$ is complete for the norm.

Answer:

(a) By Taylor's theorem, we have

$$f(x) = f(0) + \sum_{i=1}^{k-1} f^{(i)}(0)\frac{x^i}{i!} + \frac{1}{(k-1)!}\int_0^x f^{(k)}(t)(x-t)^{k-1}dt.$$

Define the function

$$r_x(t) = \sum_{i=0}^{k-1} \frac{x^i}{i!}\frac{t^i}{i!} + \frac{(-1)^k}{(2k-1)!}\max\{x-t,0\}^{2k-1} + \sum_{i=0}^{k-1}(-1)^{k+i+1}\frac{x^{2k-1-i}}{(2k-1-i)!}\frac{t^i}{i!}.$$

Then

$$r_x^{(i)}(0) = \frac{1}{i!}x^i + \frac{(-1)^{k+i}}{(2k-i-1)!}\max\{x,0\}^{2k-1-i} + \frac{(-1)^{k+i+1}}{(2k-1-i)!}x^{2k-1-i} = x^i$$

for $i < k$ and

$$r_x^{(k)}(t) = \frac{1}{(k-1)!}\max\{x-t,0\}^{k-1}.$$

Thus we have

$$\begin{aligned}
\langle f, r_x \rangle &= f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \cdots + \frac{1}{(k-1)!}f^{(k-1)}(0)x^{k-1} + \frac{1}{(k-1)!}\int_0^1 f^{(k)}(t)\,[x-t]_+^{k-1}\,dt \\
&= \sum_{i=0}^{k-1}\frac{f^{(i)}(0)}{i!}x^i + \frac{1}{(k-1)!}\int_0^x f^{(k)}(t)(x-t)^{k-1}dt \\
&= f(x)
\end{aligned}$$

where the last equality is Taylor's theorem.

(b) For the reproducing kernel, note that

$$\begin{aligned}
\mathsf{k}(x,z) &= \langle r_x, r_z \rangle \\
&= \sum_{i=0}^{k-1}\frac{x^i}{i!}\frac{z^i}{i!} + \frac{1}{(k-1)!(k-1)!}\int_0^1 [x-t]_+^{k-1}\,[z-t]_+^{k-1}\,dt \\
&= \sum_{i=0}^{k-1}\frac{x^i}{i!}\frac{z^i}{i!} + \frac{1}{(k-1)!(k-1)!}\int_0^{\min\{x,z\}} (x-t)^{k-1}(z-t)^{k-1}dt.
\end{aligned}$$

(c) To see that $\mathcal{F}_k$ is a Hilbert space, we must show that $\|f\|_{\mathcal{H}}^2 = \langle f, f\rangle$ is a norm and that $\mathcal{F}_k$ is complete for $\|\cdot\|_{\mathcal{H}}$. Non-negativity of $\|\cdot\|_{\mathcal{H}}$ and the triangle inequality are trivial, as it is clear that $\langle \cdot, \cdot\rangle$ is an inner product. Now suppose that $\|f\|_{\mathcal{H}} = 0$. Then $f^{(l)}(0) = 0$ for all $l < k$, and $\int_0^1 f^{(k)}(t)^2 dt = 0$, so that $f^{(k)} = 0$ almost everywhere. Of course, this shows that $f^{(k-1)} \equiv 0$ by integration, and so on, so that $f \equiv 0$. To show completeness, let f_n be a Cauchy sequence in $\mathcal{F}_k$. Then since

$$\|f_n - f_m\|_{\mathcal{H}}^2 = \sum_{l=0}^{k-1} (f_n^{(l)}(0) - f_m^{(l)}(0))^2 + \int_0^1 (f_n^{(k)}(t) - f_m^{(k)}(t))^2 dt,$$

it is clear that $f_n^{(l)}(0)$ is a Cauchy sequence in $\mathbb{R}$ and $f_n^{(k)}$ is a Cauchy sequence in $L^2([0,1])$. Completeness of $\mathbb{R}$ and completeness of L^2 then imply the existence of $\lim_n f_n^{(l)}(0)$ for $l < k$ and a $g \in L^2([0,1])$ such that $f_n^{(k)} \to g$ in L_2. Now define the functions $f^{(l)}$ by

$$f^{(k)}(x) = g(x), \quad f^{(k-1)}(x) = \lim_n f_n^{(k-1)}(0) + \int_0^x g(t)dt, \quad \ldots, \quad f(x) = \lim_n f_n(0) + \int_0^x f^{(1)}(t)dt.$$

Since $f^{(k)} \in L^2([0,1])$, it is clear that each of the $f^{(l)}$ are absolutely continuous, and the derivative of $f^{(l)}$ is $f^{(l+1)}$. So f_n indeed has a limit f.

□

Question 4: The variation distance between probability distributions P and Q on a space $\mathcal{X}$ is defined by $\|P - Q\|_{\text{TV}} = \sup_{A \subset \mathcal{X}} |P(A) - Q(A)|$.

(a) Show that

$$2\|P - Q\|_{\text{TV}} = \sup_{f : \|f\|_\infty \le 1} \{\mathbb{E}_P[f(X)] - \mathbb{E}_Q[f(X)]\}$$

where the supremum is taken over all functions with $f(x) \in [-1, 1]$, and the first expectation is taken with respect to P and the second with respect to Q. You may assume that P and Q have densities.

Answer: Using the assumption that we have a density and that $P(A) - Q(A) = 1 - P(A^c) - (1 - Q(A^c)) = Q(A^c) - P(A^c)$, we have

$$\begin{aligned}\|P - Q\|_{\text{TV}} = \sup_{A \subset \mathcal{X}} \{P(A) - Q(A)\} &= \sup_A \int \mathbf{1}\{x \in A\}(p(x) - q(x))dx \\ &= \int \mathbf{1}\{p(x) \ge q(x)\}(p(x) - q(x))dx.\end{aligned}$$

Similarly, we have $\|P - Q\|_{\text{TV}} = \sup_A \{Q(A) - P(A)\}$, and combining these yields

$$2\|P - Q\|_{\text{TV}} = \int (\mathbf{1}\{p(x) \ge q(x)\} - \mathbf{1}\{p(x) \le q(x)\})(p(x) - q(x))dx.$$

But of course, $\sup_{a \in [-1,1]} a(p - q) = (p - q)(\mathbf{1}\{p \ge q\} - \mathbf{1}\{p \le q\})$, which proves the result. □

Question 5: In a number of experimental situations, it is valuable to determine if two distributions P and Q are the same or different. For example, P may be the distribution of widgets produced by one machine, Q the distributions of widgets by a second machine, and we wish to test if the two distributions are the same (to within allowable tolerances). Let $\mathcal{H}$ be an RKHS of functions with domain $\mathcal{X}$ and reproducing kernel k, and let P and Q be distributions on $\mathcal{X}$.

(a) Let $\|\cdot\|_{\mathcal{H}}$ denote the norm on the Hilbert space $\mathcal{H}$. Show that

$$D_{\mathsf{k}}(P,Q)^2 := \sup_{f:\|f\|_{\mathcal{H}}\le 1}\left\{|\mathbb{E}_P[f(X)] - \mathbb{E}_Q[f(Z)]|^2\right\} = \mathbb{E}[\mathsf{k}(X,X')] + \mathbb{E}[\mathsf{k}(Z,Z')] - 2\mathbb{E}[\mathsf{k}(X,Z)]$$

where $X, X' \stackrel{\text{iid}}{\sim} P$ and $Z, Z' \stackrel{\text{iid}}{\sim} Q$.

(b) A kernel $\mathsf{k} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is called *universal* if the induced RKHS $\mathcal{H}$ of functions $f : \mathcal{X} \to \mathbb{R}$ can arbitrarily approximate continuous functions. That is, for any $\phi : \mathcal{X} \to \mathbb{R}$ continuous and $\epsilon > 0$, there is some $f \in \mathcal{H}$ such that

$$\sup_{x\in\mathcal{X}} |f(x) - \phi(x)| \le \epsilon.$$

Show that if k is universal, then

$$D_{\mathsf{k}}(P,Q) = 0 \ \text{ if and only if } \ P = Q.$$

You may assume $\mathcal{X}$ is a metric space and that $P = Q$ iff $P(A) = Q(A)$ for all compact $A \subset \mathcal{X}$.

(c) You wish to estimate $D_{\mathsf{k}}(P,Q)$ given samples from each of the distributions. Assume that $\mathsf{k}(x,z) \in [-B,B]$ for all $x, z \in \mathcal{X}$. Let $X_i \stackrel{\text{iid}}{\sim} P$, $i = 1,\ldots,n_1$ and $Z_i \stackrel{\text{iid}}{\sim} Q$, $i = 1,\ldots,n_2$. Define

$$\widehat{K}(X_{1:n_1}) := \binom{n_1}{2}^{-1} \sum_{1\le i<j\le n_1} \mathsf{k}(X_i,X_j), \quad \widehat{K}(Z_{1:n_2}) := \binom{n_2}{2}^{-1} \sum_{1\le i<j\le n_2} \mathsf{k}(Z_i,Z_j),$$

and

$$\widehat{K}(X_{1:n_1}, Z_{1:n_2}) := \frac{1}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \mathsf{k}(X_i,Z_j).$$

Show that $\mathbb{E}[\widehat{K}(X_{1:n})] = \mathbb{E}[\mathsf{k}(X,X')]$ and $\mathbb{E}[\widehat{K}(X_{1:n_1}, Z_{1:n_2})] = \mathbb{E}[\mathsf{k}(X,Z)]$ for $X, X' \stackrel{\text{iid}}{\sim} P$ and $Z, Z' \stackrel{\text{iid}}{\sim} Q$. Show for some numerical constant $c > 0$ that for all $t \ge 0$,

$$\mathbb{P}\left(\left|\widehat{K}(X_{1:n}) - \mathbb{E}[\mathsf{k}(X,X')]\right| \ge t\right) \le 2\exp\left(-c\frac{nt^2}{B^2}\right)$$

and

$$\mathbb{P}\left(\left|\widehat{K}(X_{1:n_1}, Z_{1:n_2}) - \mathbb{E}[\mathsf{k}(X,Z)]\right| \ge t\right) \le 2\exp\left(-c\frac{n_1 t^2}{B^2}\right) + 2\exp\left(-c\frac{n_2 t^2}{B^2}\right).$$

(d) Define the empirical Hilbert distances

$$\widehat{D}^2_{\mathsf{k}}(P,Q) := \binom{n_1}{2}^{-1} \sum_{1\le i<j\le n_1} \mathsf{k}(X_i,X_j) + \binom{n_2}{2}^{-1} \sum_{1\le i<j\le n_2} \mathsf{k}(Z_i,Z_j) - \frac{2}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \mathsf{k}(X_i,Z_j).$$

Show that for all $t \ge 0$,

$$\mathbb{P}\left(\left|\widehat{D}^2_{\mathsf{k}}(P,Q) - D^2_{\mathsf{k}}(P,Q)\right| \ge t\right) \le C\exp\left(-c\frac{\min\{n_1,n_2\}t^2}{B^2}\right)$$

where $0 < c$, $C < \infty$ are numerical constants.

Answer:

(a) As $\mathsf{k} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the reproducing kernel for $\mathcal{H}$, we have for any $f \in \mathcal{H}$ such that $\|f\|_{\mathcal{H}} \le 1$

$$\begin{aligned}\mathbb{E}[f(X)] - \mathbb{E}[f(Z)] &= \mathbb{E}[\langle f, \mathsf{k}(X, \cdot)\rangle] - \mathbb{E}[\langle f, \mathsf{k}(Z, \cdot)\rangle] \\ &\stackrel{(i)}{=} \langle f, \mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]\rangle \\ &\stackrel{(ii)}{\le} \|f\|_{\mathcal{H}} \|\mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]\|_{\mathcal{H}} \le \|\mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]\|_{\mathcal{H}},\end{aligned}$$

where we have used linearity in (i) and Cauchy-Schwarz in (ii), and that $\|f\|_{\mathcal{H}} \le 1$ in the final line. Equality holds in step (ii) if

$$f(\cdot) = \frac{\mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]}{\|\mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]\|_{\mathcal{H}}},$$

and we have

$$\begin{aligned}\|\mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)]\|_{\mathcal{H}}^2 &= \left\langle \mathbb{E}[\mathsf{k}(X, \cdot) - \mathsf{k}(Z, \cdot)], \mathbb{E}[\mathsf{k}(X', \cdot) - \mathsf{k}(Z', \cdot)]\right\rangle \\ &= \left\langle \mathbb{E}[\mathsf{k}(X, \cdot)], \mathbb{E}[\mathsf{k}(X', \cdot)]\right\rangle + \left\langle \mathbb{E}[\mathsf{k}(Z, \cdot)], \mathbb{E}[\mathsf{k}(Z', \cdot)]\right\rangle - 2\left\langle \mathbb{E}[\mathsf{k}(X, \cdot)], \mathbb{E}[\mathsf{k}(Z, \cdot)]\right\rangle \\ &= \mathbb{E}[\mathsf{k}(X, X')] + \mathbb{E}[\mathsf{k}(Z, Z')] - 2\mathbb{E}[\mathsf{k}(X, Z)],\end{aligned}$$

where the final equality uses the linearity of the inner product and independence of X, X', Z, Z'.

(b) Suppose that $P = Q$. Then certainly $\mathbb{E}_P[f(X)] - \mathbb{E}_Q[f(Z)] = \mathbb{E}_P[f(X)] - \mathbb{E}_P[f(X)] = 0$ for all $f \in \mathcal{H}$. Now suppose $P \neq Q$. Then there exists a compact set A such that $P(A) \neq Q(A)$. For $n \in \mathbb{N}$, define the function

$$\phi_n(x) = \max\{1 - n \cdot \operatorname{dist}(x, A), 0\} = [1 - n \operatorname{dist}(x, A)]_+,$$

which satisfies $\phi_n(x) = 1$ for $x \in A$, $\phi_n(x) = 0$ for x such that $\operatorname{dist}(x, A) \ge 1/n$, and is Lipschitz continuous. Moreover, we have $\phi_n(x) \downarrow \mathbf{1}\{x \in A\}$ for all $x \in A$ as $n \to \infty$. Thus the monotone convergence theorem gives that

$$\lim_n \mathbb{E}_P[\phi_n(X)] = P(A) \quad \text{and} \quad \lim_n \mathbb{E}_Q[\phi_n(Z)] = Q(A).$$

Let $\epsilon > 0$ be such that $|P(A) - Q(A)| \ge 4\epsilon$. Choose N such that $n \ge N$ implies $|\mathbb{E}_P[\phi_n] - P(A)| < \epsilon$ and $|\mathbb{E}_Q[\phi_n] - Q(A)| < \epsilon$, and let $n \ge N$. Choose $f \in \mathcal{H}$ such that $\sup_x |f(x) - \phi_n(x)| \le \epsilon$. Then

$$|\mathbb{E}_P[f(X)] - \mathbb{E}_Q[f(Z)]| \ge |\mathbb{E}_P[\phi_n(X)] - \mathbb{E}_Q[\phi_n(Z)]| - 2\epsilon > |P(A) - Q(A)| - 4\epsilon \ge 4\epsilon - 4\epsilon = 0.$$

Dividing by $\|f\|_{\mathcal{H}}$ we have

$$D_{\mathsf{k}}(P, Q) = \sup_{g : \|g\|_{\mathcal{H}} \le 1} |\mathbb{E}_P[g] - \mathbb{E}_Q[g]| \ge \frac{|\mathbb{E}_P[f(X)] - \mathbb{E}_Q[f(Z)]|}{\|f\|_{\mathcal{H}}} > 0.$$

(c) The expectation equalities are immediate.

We apply bounded differences for the first statement. We first look at $f(x_{1:n}) = \widehat{K}(x_{1:n})$. As the function is symmetric, we fix index $i = 1$. Then for $x, x' \in \mathcal{X}$, we have

$$f(x, x_{2:n}) - f(x', x_{2:n}) = \binom{n}{2}^{-1} \sum_{j=2}^{n} (\mathsf{k}(x, X_j) - \mathsf{k}(x', X_j))$$

and using that $\mathsf{k}(x, x') \in [-B, B]$, the summands are each bounded by $2B$ in magnitude. Thus

$$|f(x, x_{2:n}) - f(x', x_{2:n})| \le \frac{2}{n(n-1)} \cdot 2B(n-1) = \frac{4B}{n}.$$

Bounded differences (McDiarmid's inequality) implies

$$\mathbb{P}\left(\left|\widehat{K}(X_{1:n}) - \mathbb{E}[\widehat{K}(X_{1:n})]\right| \ge t\right) \le 2\exp\left(-\frac{nt^2}{8B^2}\right).$$

The argument about $\widehat{K}(X_{1:n_1}, Z_{1:n_2})$ is a bit more complex. Define

$$\widehat{K}(X_{1:n_1}, Q) = \frac{1}{n_1}\sum_{i=1}^{n_1} \mathbb{E}_Q[\mathsf{k}(X_i, Z) \mid X_i].$$

Then we have

$$\mathbb{E}[\widehat{K}(X_{1:n_1}, Z_{1:n_2}) \mid X_{1:n_1}] = \widehat{K}(X_{1:n_1}, Q)$$

by the independence of Z_i, X_j. Fixing $X_{1:n_1}$, define the function $g(z_{1:n_2} \mid X_{1:n_1})$ by

$$g(z_{1:n_2} \mid X_{1:n_1}) = \widehat{K}(X_{1:n_1}, z_{1:n_2}).$$

Then g satisfies bounded differences with parameter $4B/n_2$, as above, and so *conditional on* $X_{1:n_1}$, we have

$$\mathbb{P}\left(\left|g(Z_{1:n_2} \mid X_{1:n_1}) - \widehat{K}(X_{1:n_1}, Q)\right| \ge t \mid X_{1:n_1}\right) \le 2\exp\left(-\frac{n_2 t^2}{8B^2}\right). \tag{1}$$

Now we argue that

$$x_{1:n_1} \mapsto \widehat{K}(x_{1:n_1}, Q)$$

satisfies bounded differences as well. Note that $\mathbb{E}[\widehat{K}(X_{1:n_1}, Q)] = \mathbb{E}[\mathsf{k}(X, Z)]$ by construction. Without loss of generality let us fix $x_{2:n_1}$ and modify $x_1 \in \{x, x'\}$. Then

$$\widehat{K}(x, x_{2:n_1}, Q) - \widehat{K}(x', x_{2:n_1}, Q) = \frac{1}{n_1}\mathbb{E}_Q[\mathsf{k}(x, Z) - \mathsf{k}(x', Z)] \in \left[-\frac{2B}{n_1}, \frac{2B}{n_1}\right],$$

satisfying bounded differences with parameter $2B/n_1$. Thus we have

$$\mathbb{P}\left(\left|\widehat{K}(X_{1:n_1}, Q) - \mathbb{E}[\mathsf{k}(X, Z)]\right| \ge t\right) \le 2\exp\left(-\frac{n_1 t^2}{2B^2}\right). \tag{2}$$

Combining the bounds (1) and (2) and applying the tower property of expectation and the triangle inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(\left|\widehat{K}(X_{1:n_1}, Z_{1:n_2}) - \mathbb{E}[\mathsf{k}(X, Z)]\right| \ge t\right) \\
&\le \mathbb{E}\left[\mathbb{P}\left(\left|g(Z_{1:n_2} \mid X_{1:n_1}) - \widehat{K}(X_{1:n_1}, Q)\right| \ge t/2 \mid X_{1:n_1}\right)\right] + \mathbb{P}\left(\left|\widehat{K}(X_{1:n_1}, Q) - \mathbb{E}[\mathsf{k}(X, Z)]\right| \ge t/2\right) \\
&\le 2\exp\left(-\frac{n_2 t^2}{32B^2}\right) + 2\exp\left(-\frac{n_1 t^2}{8B^2}\right).
\end{aligned}$$

$\square$