1 Reproducing kernel Hilbert spaces

(a) Let $\mathcal{H}$ be a RKHS generated by a kernel $K$. Consider training data $(X_1, Y_1), \ldots, (X_n, Y_n)$ where $Y_i \in \{-1, +1\}$. Let $\lambda \geq 0$ be a fixed real number, and let $\hat{f}$ minimize $\sum_{i=1}^{n} L(f(X_i), Y_i) + \lambda \|f\|^2_K$. Show that $\hat{f}$ has the form

$$\hat{f}(x) = \sum_{i=1}^{n} \hat{\alpha}_i K(X_i, X)$$

where $\hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_n)$ minimizes

$$Q(\kappa \alpha) + \lambda \alpha^T \kappa \alpha$$

for some function $Q$ where $\kappa$ is an $n \times n$ matrix with $\kappa(j, k) = K(X_j, X_k)$.

(b) Let $\mathcal{F}$ be the set of all functions $f : [0, 1] \to \mathbb{R}$ such that $f(x) = ax$ for some real number $a$. Show that this is a RKHS with kernel $K(x, y) = xy$.

2 Basic inequalities

(a) Show that there exists a random variable $X \geq 0$ such that Markov’s inequality is an equality.

(b) Show that there exists a random variable $Y$ such that Chebyshev’s inequality is an equality.

3 Chernoff’s method

Recall Chernoff’s method where we use the fact that $P(X > \delta) \leq \inf_{t > 0} e^{-t\delta} E[e^{tX}]$. Let $X \geq 0$. Suppose that the moment generating function for $X$ exists. Let $\delta > 0$. Show that

$$\inf_{k=0,1,2,\ldots} \frac{E[X^k]}{\delta^k} \leq \inf_{t > 0} e^{-t\delta} E[e^{tX}]$$

(Hence, Chernoff’s method does not necessarily give the tightest possible bounds.)

4 Sub-Gaussian variables

A random variable is sub-Gaussian if there exists $a > 0$ such that for all $t \in \mathbb{R}$

$$E[e^{t(X-\mu)}] \leq e^{\frac{a^2t^2}{2}}.$$
Let $X$ be Rademacher. Thus $P(X = 1) = P(X = -1) = 1/2$. Show that $X$ is sub-Gaussian with $a = 1$.

(b) Show that every bounded random variable is sub-Gaussian.

(c) Show that, if $X$ is sub-Gaussian then
$$P(|X - \mu| \geq t) \leq 2e^{-\frac{t^2}{2a^2}}.$$ for all $t \geq 0$, where $Z \sim N(0, b^2)$.

You may use the following fact without proof: Let $Z \sim N(0, 1)$ and $z > 0$. Let $\phi(z)$ be the standard Normal density. Then $P(Z \geq z) \geq \phi(z)(1/z - 1/z^3)$.

Hint: Consider the ratio $P(X \geq t)/P(Z \geq t)$. Now consider two cases: (i) $0 \leq t \leq 2\sigma$ and (ii) $t > 2\sigma$.

5 Inequality on a convex function

Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex, non-decreasing function. Let $\mathcal{F}$ be a class of functions. Let $\mathcal{G} = \{f - \mathbb{E}[f(X)] : f \in \mathcal{F}\}$. Let \[
\xi_n = \sup_{f \in \mathcal{G}} \frac{1}{n} \sum_{i=1}^{n} \sigma_i f(X_i)
\]
where $\sigma_1, \ldots, \sigma_n$ are Rademacher. Show that \[
\mathbb{E}\left[\phi\left(\sup_{f \in \mathcal{F}} |P_n(f) - P(f)|\right)\right] \geq \mathbb{E}\left[\phi\left(\frac{1}{2n} \xi_n\right)\right].
\]

6 Rademacher complexity

Let $\mathcal{F}$ denote all indicator functions of finite subsets of $[0, 1]$. Let $P$ be the uniform distribution on $[0, 1]$. Show that $\text{Rad}_n(\mathcal{F}) \geq 1/2$, where \[
\text{Rad}_n(\mathcal{F}) = \mathbb{E}\left(\sup_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} \sigma_i f(Z_i)\right).
\]

7 Minimax rate

Let $X_1, \ldots, X_n \sim \text{Unif}(0, \theta)$ where $0 < \theta < M$ for some constant $M$. Let \[
R_n = \inf_{\theta} \sup_{\theta \in (0, M)} \mathbb{E}[\hat{\theta} - \theta]
\]
be the minimax risk for estimating $\theta$. Show that $R_n \asymp 1/n$. (Recall that $a_n \asymp b_n$ means that both $a_n/b_n$ and $b_n/a_n$ are bounded for all large $n$.)
8 Hellinger distance

The Hellinger distances between distributions $P$ and $Q$ is

$$H(P, Q) = \sqrt{\int \left(\sqrt{p} - \sqrt{q}\right)^2}$$

where $P$ has density $p$ and $Q$ has density $q$.

(a) Let $TV$ denote the total variation distance. Show that

$$TV(P, Q) \leq H(P, Q) \sqrt{1 - \frac{H^2(P, Q)}{4}}.$$  

You may use the following fact without proof: $TV(P, Q) = \|P - Q\|_1 / 2$.

(b) Let $X_1, \ldots, X_n$ be iid. Let $p^n = p(x_1, \ldots, x_n) = \prod_i p(x_i)$ and $q^n = q(x_1, \ldots, x_n) = \prod_i q(x_i)$ be joint density functions. Let $p_i = p(x_i)$ and $q_i = q(x_i)$. Show that

$$H(p^n, q^n) = \sqrt{2} \sqrt{1 - \prod_{i=1}^{n} (1 - (1/2)H^2(p_i, q_i))}.$$