Theoretical Machine Learning

Lectured by Zhihua Zhang LaTeXed by Chengxin Gong

March 25, 2024

Contents

1 简介

- 机器学习的主要任务: 生成、预测、决策. 生成: $X_1, \cdots, X_n \sim F$, 推断分析 F, 无监督学习, GAN, GPT, $\cdots$. 预测: 数据对 $(X^{(1)}, Y^{(1)}), \cdots, (X^{(n)}, Y^{(n)})$, $X^{(i)} \in \mathbb{R}^d$ 输入变量, $f: \mathcal{X} \to \mathcal{Y}, x \in \mathcal{X}, y \in \mathcal{Y}$, 归因, 有监督学习. 决策: 强化学习, Agent←action, state, reward→ 环境.

- 求解问题的途径: 参数/非参数, 频率 (MLE)/贝叶斯.

- 误差模型: 有监督: $X = (X_1, \cdots, X_d)^T \in \mathbb{R}^d$, 回归: $Y \in \mathbb{R}$; 分类: $Y \in \{0,1\}(\{-1,1\}, \{1, \cdots, M\}, \{0,1\}^M)$; X 随机, Random design(生成模型), $Y = g(X) + \varepsilon \overset{\text{or}}{=} g(X, Z), Y^{(i)} = g(X^{(i)}, Z^{(i)})$; X 固定 $X = x$, Fixed design(判别模型), $Y^{(i)} = g(x^{(i)}, Z^{(i)})$. 无监督: $X = g(Z)$(因子模型: $X = AZ + \varepsilon, Z \in \mathcal{N}(0,1), \varepsilon \sim \mathcal{N}(0, \Sigma)$).

2 统计决策理论

- Consider a state space Ω, data space $\mathcal{D}$, model $\mathcal{P} = \{p(\theta, x)\}$, action space $\mathscr{A}$. Loss function: $\mathcal{L}: \Omega \times \mathscr{A} \to [-\infty, +\infty]$, measurable, nonnegative. A measurable function $\delta: \mathcal{D} \to \mathscr{A}$ is called a nonrandomized decision rule. Risk function is defined as $\mathcal{R}(\theta, \delta) = \int \mathcal{L}(\theta, \delta(x)) \mathrm{d}P_\theta(x) = \mathbb{E}_\theta \mathcal{L}(\theta, \delta(X))$. Randomized decision: for each $X = x$, $\delta(x)$ is a probability distribution: $[A|X = x] \sim \delta_x$. Risk function for δ: $\mathcal{R}(\theta, \delta) = \mathbb{E}_\theta \mathcal{L}(\theta, A) = \mathbb{E}_\theta \mathbb{E}_a \mathcal{L}(\theta, A|X) = \iint \mathcal{L}(\theta, a) \mathrm{d}\delta_x(a) \mathrm{d}P_\theta(x)$.

- Example [参数估计]: $\theta \in \Omega, \mathscr{A} = \Omega, \mathcal{L}(\theta, a) = \|\theta - a\|_2^2 \overset{\text{or}}{=} \|\theta - a\|_p^p (p \geq 1) \overset{\text{or}}{=} \int \log \frac{P_\theta(x)}{P_a(x)} P_\theta(x) \mathrm{d}m(x) (\text{KL})$. $\mathcal{R} = \mathrm{Var}(a) + \mathrm{bias}^2(a)$. Bregmass loss: $\phi: \mathbb{R}^d \to \mathbb{R}$ describe any strictly convex differentiable function. Then $\mathcal{L}_\phi(\theta, a) = \phi(a) - \phi(\theta) - (\phi - a)^T \nabla \phi(a)$.

- Example [Testing]: $\mathscr{A} = \{0, 1\}$ with action "0" associated with accepting $H_0: \theta \in \Omega_0$ and "1": $H_1: \theta \in \Omega_1$. δ_x is a Bernolli distribution. $\mathcal{L}(\theta, a) = I\{a = 1, \theta \in \Omega_0\} + I\{a = 0, \theta \in \Omega_1\}$. Risk $\mathcal{R}(\theta, \delta) = \mathbb{P}_\theta(A = 1) 1_{\theta \in \Omega_0} + \mathbb{P}_\theta(A = 0) 1_{\theta \in \Omega_1}$.

- A decision rule δ is called inadmissible if a competing rule δ^* such that $\mathcal{R}(\theta, \delta^*) \leq \mathcal{R}(\theta, \delta)$ for all $\theta \in \Omega$ and $\mathcal{R}(\theta, \delta^*) < \mathcal{R}(\theta, \delta)$ for at least one $\theta \in \Omega$. Otherwise, δ is admissible.

- The maximum risk $\bar{\mathcal{R}}(\delta) = \sup_{\theta \in \Omega} \mathcal{R}(\theta, \delta)$ and the Bayes risk $r(\Lambda, \delta) = \int \mathcal{R}(\theta, \delta) \mathrm{d}\Lambda(\theta) = \int \mathcal{L}(\theta, \delta) \mathrm{d}\mathbb{P}(x, \theta)$ ($\Lambda(\theta)$ is a prior). A decision rule that minimizes the Bayes risk is called a Bayes rule, that is, $\hat{\delta}: r(\Lambda, \hat{\delta}) = \inf_\delta r(\Lambda, \delta)$. Minimax rule $\delta^*: \sup_{\theta \in \Omega} \mathcal{R}(\theta, \delta^*) = \inf_\delta \sup_{\theta \in \Omega} \mathcal{R}(\theta, \delta)$.

- If risk functions for all decision rules are continuous in θ, if δ is Bayesian for Λ and has finite integrated risk $r(\Lambda, \delta) < \infty$, and if the support of Λ is the whole state space Ω, then δ is admissible.

- $p(\theta|x) = \frac{p_\theta(x)\lambda(\theta)}{\int p_\theta(x)\lambda(\theta)\mathrm{d}\theta} := \frac{p_\theta(x)\lambda(\theta)}{m(x)}$. Define the posterior risk of δ: $r(\delta|X = x) = \int \mathcal{L}(\theta, \delta(x)) \mathrm{d}\mathbb{P}(\theta|x)$. The Bayes risk $r(\Lambda, \delta)$ satisfies that $r(\Lambda, \delta) = \int r(\delta|x) \mathrm{d}M(x)$. Let $\hat{\delta}(x)$ be the value of δ that minimizes $r(\delta|x)$. Then $\hat{\delta}$ is the Bayes rule.

- Application to supervised learning. Case 1: Regression. $(X, Y) \in \mathcal{X} \times \mathcal{Y}, f: \mathcal{X} \to \mathcal{Y}, \mathscr{A} = \Omega = \mathcal{Y}, \mathcal{D} = \mathcal{X}, \delta = f, \mathcal{L}(Y, f(X)) = \|Y - f(X)\|_p^p, p \geq 1$, risk $R_f = \iint \mathcal{L}(y, f(x)) \mathrm{d}\mathbb{P}(x, y) = \mathbb{E}[\mathcal{L}(Y, f(X))] = \mathbb{E}[\mathbb{E}\mathcal{L}(Y, f(X))|X]$. When $p = 2$, $r(f|X = x) = \int \mathcal{L}(y, f(x)) \mathrm{d}\mathbb{P}(y|x) = \int |y - f(x)|^2 \mathrm{d}\mathbb{P}(y|x)$. 回归函数 $g(x) := \int y \mathrm{d}\mathbb{P}(y|x) \Rightarrow R_f = \mathbb{E}|Y - f(X)|^2 = \mathbb{E}|Y - g(X) + g(X) - f(X)|^2 = \mathbb{E}|Y - g(X)|^2 + \mathbb{E}|g(X) - f(X)|^2 \geq \mathbb{E}|Y - g(X)|^2$.

- Case 2: Pattern classification. $Y \in \{0, 1\}, p_0 = P(Y = 0), p_1 = \mathbb{P}(Y = 1) = 1 - p_0, \mathbb{E}[\mathcal{L}(Y, f(X))] = \mathbb{P}(Y \neq f(X))$. The Bayesian rule (predictor) is given by $f(x) = 1\{\mathbb{P}(Y = 1|X = x) \geq \frac{\mathcal{L}(1,0) - \mathcal{L}(0,0)}{\mathcal{L}(0,1) - \mathcal{L}(1,1)} \mathbb{P}(Y = 0|X = x)\}$. (Proof: $\mathbb{E}[\mathcal{L}(Y, f(X))|X = x] = \begin{cases} \mathbb{E}[\mathcal{L}(Y, 0)|X = x] = \mathcal{L}(0,0)\mathbb{P}(Y = 0|X = x) + \mathcal{L}(1,0)\mathbb{P}(Y = 1|X = x) \\ \mathbb{E}[\mathcal{L}(Y, 1)|X = x] = \mathcal{L}(0,1)\mathbb{P}(Y = 0|X = x) + \mathcal{L}(1,1)\mathbb{P}(Y = 1|X = x) \end{cases}$, 比较大小)

- 连续化: $\mathbb{P}(Y = 1|X = x) = \mathbb{E}(Y|X = x) := g(x)$(回归), $f(x) = 1\{g(x) \geq \frac{1}{2}\}$. Then $0 \leq \mathbb{P}(\hat{f}(X) \neq Y) - \mathbb{P}(f(X) \neq Y) \leq 2\int_{\mathcal{X}} |\hat{g}(x) - g(x)| \mu(\mathrm{d}x) \leq 2(\int_{\mathcal{X}} |\hat{g}(x) - g(x)|^2 \mu(dx))^{\frac{1}{2}}$.

- 回到 Case 2. $f(x)=1\{\frac{p(x|y=1)}{p(x|y=0)}\geq\frac{p_0(\mathcal{L}(0,1)-\mathcal{L}(0,0))}{p_1(\mathcal{L}(1,0)-\mathcal{L}(1,1))}\}$, 这与似然比检验 (LRT) 相同: Likelihood $L(X):=\frac{p(X|Y=1)}{p(X|Y=0)}$, 形式为 $f(x)=1\{L(x)\geq\eta\}$.

- Confusion table:

	$Y=0$	$Y=1$
$\hat{Y}=0$	true negative	false negative
$\hat{Y}=1$	false positive	true positive

  Ture Positive Rate: TPR $=\mathbb{P}(\hat{Y}=1|Y=1)$; False Negative Rate: FNR = 1 - TPR, type II error; False Positive Rate: FPR $=\mathbb{P}(\hat{Y}=1|Y=0)$, type I error; True Negative Rate: TNR = 1 - FPR. Precision: $\mathbb{P}(Y=1|\hat{Y}=1)=\frac{p_1\text{TPR}}{p_0\text{FPR}+p_1\text{TPR}}$. F_1-score: F_1 is the harmonic mean of precision and recall, which can be written as $F_1=\frac{2\text{TPR}}{1+\text{TPR}+\frac{p_0}{p_1}\text{FPR}}$.

- Optimization: maximize TPR subject to FPR $\leq\alpha,\alpha\in[0,1]$. Randomized rule: Q return 1 with probability $Q(x)$ and 0 with probability $1-Q(x)$. Maximize $\mathbb{E}[Q(x)|Y=1]$ subject to $\mathbb{E}[Q(x)|Y=0]\leq\alpha$. Suppose the likelihood functions $p(x|y)$ are continuous. Then the optimal predictor is a deterministic LRT (N-P lemma). (Proof: Let η be the threshold for an LRT such that the predictor $Q_\eta(x)=1\{\alpha(x)\geq\eta\}$ has FPR $=\alpha$. Such an LRT exists because likelihood are continuous. Let β denote the TPR of Q_η. Prove that Q_η is optimal for risk minimization problem corresponding to the loss functions $\mathcal{L}(0,1)=\eta\frac{p_1}{p_0},\mathcal{L}(1,0)=1,\mathcal{L}(1,1)=\mathcal{L}(0,0)=0$ since $\frac{p_0(\mathcal{L}(0,1)-\mathcal{L}(0,0))}{p_1(\mathcal{L}(1,0)-\mathcal{L}(1,1))}=\frac{p_0\mathcal{L}(0,1)}{p_1\mathcal{L}(1,0)}=\eta$. Under these loss functions, the risk of Bayes predictor for Q is $\mathcal{R}_Q=p_0\text{FPR}(Q)\mathcal{L}(0,1)+p_1(1-\text{TPR}(Q))\mathcal{L}(1,0)=p_1\eta\text{FPR}(Q)+p_1(1-\text{TPR}(Q))$. Now let Q be any other rule with $\text{FPR}(Q)\leq\alpha$, $\mathcal{R}_{Q_\eta}=p_1\eta\alpha+p_1(1-\beta)\leq p_1\eta\text{FPR}(Q)+p_1(1-\text{TPR}(Q))\leq p_1\eta\alpha+p_1(1-\text{TPR}(Q))\Rightarrow\text{TPR}(Q)\leq\beta$)

- ROC (Receiver operating character) curve: y-axis is TPR and x-axis is FPR. Proposition: (1) The points $(0,0)$ and $(1,1)$ are on the ROC curve; (2) The ROC must lie above the main diagnal; (3) The ROC curve is concave. (Proof: (2): Fix $\alpha\in(0,1)$ and consider a randomized rate TPR = FPR $=\alpha$, $Q(x)\equiv\alpha$; (3): Consider two rules (FPR(η_1), TPR(η_1)) and (FPR(η_2), TPR(η_2)). If we flip a biased coin and use the first rule with probability t and use the second rule with probability $1-t$. Then this yields a randomized rule with (FPR, TPR) = (tFPR(η_1) + $(1-t)$FPR(η_2), tTPR(η_1) + $(1-t)$FPR(η_2)). Fixing FPR $\leq t$FPR(η_1) + $(1-t)$FPR(η_2), TPR $\geq t$TPR(η_1) + $(1-t)$TPR(η_2).)

- Markov Decision Processes (MDPs): Five elements: decision epoches, states, actions, transition probabilities and rewards. (1) Decision epoches: Let T denote the set of decision epoches, discrete: $\{1,2,\cdots,N\}$; continuous: $[0,N]$; $N</=\infty$: finite or infinite. (2) State and action sets: decision epoch $t\in T$, the system occupies a state $S_t\in\mathcal{S}$, the decision maker $a\in\mathcal{A}$. (3) Reward and transition probabilities: t, in state s, choose action a, (i) the decision maker receives a reward $r_t(s,a)$, (ii) the system state at the next decision epoch is determined by the probability distribution $p_t(\cdot|s_t,a)$.

- Decision rules: Prescribe a procedure for action selection in each state at a specified decision epoch. Four cases: (1) Markovian and Deterministic: $\delta_t:\mathcal{S}\to\mathcal{A}$; (2) M and Randomized: $\delta_t:\mathcal{S}\to\Delta(\mathcal{A})(q_{\delta_t(s)}(a))$; (3) History-dependent and D: $h_t=(s_1,a_1,\cdots,s_{t-1},a_{t-1},s_t)=(h_{t-1},a_{t-1},s_t),\mathcal{H}_1=\mathcal{S},\mathcal{H}_2=\mathcal{S}\times\mathcal{A}\times\mathcal{S},\cdots,\delta_t:\mathcal{H}_t\to\mathcal{A}$; (4) HR: $\delta_t:\mathcal{H}_t\times\Delta(\mathcal{A})$. A policy $\pi=(\delta_1,\delta_2,\cdots,\delta_{N-1})$ is stationary if $\delta_1=\delta_2=\cdots=\delta$ for $t\in T$.

- Let $\pi=(\delta_1,\cdots,\delta_{N-1})$ in HR and $R_t:=r_t(X_t,Y_t)$ denote the random reward, $R_N:=r_N(X_N)$, $R:=(R_1,\cdots,R_N)$. The expected total reward $U_N^\pi(s):=\mathbb{E}^\pi\{\sum\limits_{t=1}^{N-1}r_t(X_t,Y_t)+r_N(X_N)|X_1=s\}$. Assume $|r_t(s,a)|\leq M<\infty$ for all $(s,a)\in\mathcal{S}\times\mathcal{A}$. Optimal policy: $U_N^{\pi^*}(s)\geq U_N^\pi(s),s\in\mathcal{S}$. ε-optimal policy: $U_N^{\pi_\varepsilon^*}(s)+\varepsilon>U_N^\pi(s),s\in\mathcal{S}$. The value of the MDP: $U_N^*(s)=\sup_{\pi\in\mathcal{D}^{\text{HR}}}U_N^\pi(s),s\in\mathcal{S}$.

- Finite-Horizon Policy Evaluation: $V_t^\pi(h_t)=\mathbb{E}^\pi\{\sum\limits_{k=t}^{N-1}r_k(X_k,Y_k)+r_N(X_N)|h_t\},V_N^\pi(h_N)=r_N(s),\pi\in\mathcal{D}^{\text{HD}}$. 由重期望公式, $V_t^\pi(h_t)=r_t(s_t,\delta_t(h_t))+\mathbb{E}_{h_t}^\pi V_{t+1}^\pi(h_t,\delta_t(h_t),X_{t+1})=r_t(s_t,\delta_t(h_t))+\sum\limits_{j\in\mathcal{S}}V_{t+1}^\pi(h_t,\delta_t(h_t),j)p(j|s_t,\delta_t(h_t))$.

Consider randomness (i.e. $\pi \in \mathcal{D}^{\text{HR}}$): $V_t^\pi(h_t) = \sum_{a\in\mathcal{A}} q_{\delta_t(h_t)}(a)\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} V_{t+1}^\pi(h_t,a,j)p(j|s_t,a)\}$. Computational complexity: let $K = |\mathcal{S}|, L = |\mathcal{A}|$, at decision epoch t, $K^{t+1}L^t$ histories, $K^2\sum_{i=0}^{N-1}(KL)^i$ multiplications. If $\pi \in \mathcal{D}^{\text{MD}}$, $V_t^\pi(s_t) = r_t(s_t,\delta_t(s_t)) + \sum_{j\in\mathcal{S}} V_{t+1}^\pi(j)p(j|s_t,\delta_t(s_t))$, only $(N-1)K^2$ multiplications. On the other hand, given π, this yields a valid and accurate calculation method for $U_N^\pi(s)$.

- The Bellman Equations: Let $V_t^*(h_t) = \sup_{\pi\in\mathcal{D}^{\text{HR}}} V_t^\pi(h_t)$. The optimality equations: $V_t(h_t) = \sup_{a\in\mathcal{A}}\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} V_{t+1}(h_t,a,j)p_t(j|s_t,a)\}$ for $t = 1,2,\cdots,N-1$ and $h_t = (h_{t-1},a_{t-1},s_t) \in \mathcal{H}_t$. For $t = N, V_N(h_N) = r_N(s_N)$. Suppose V_t is a solution and V_N satisfies $V_N(h_N) = r_N(s_N)$. Then $V_t(h_t) = V_t^*(h_t)$ for all $h_t \in \mathcal{H}_t, t = 1,\cdots,N$ and $V_1(s_1) = V_1^*(s_1) = U_N^*(s_1)$ for all $s_1 \in \mathcal{S}$. (Proof: Two parts. First prove $V_n(h_n) \geq V_n^*(h_n)$ for all $h_n \in \mathcal{H}_n$. By induction: $N: V_N(h_N) = r_N(s_N) = V_N^*(h_N)$ for all h_t,π. Now assume that $V_t(h_t) \geq V_t^*(h_t)$ for all $h_t \in \mathcal{H}_t$ for $t = n+1,\cdots,N$. Let $\pi' = (\delta_1',\cdots,\delta_{N-1}')$ be an arbitrary policy in $\mathcal{D}^{\text{HR}}$. For $t = n$, the Bellman equations $V_n(h_n) = \sup_{a\in\mathcal{A}}\{r_n(s_t,a_t) + \sum_{j\in\mathcal{S}} p(j|s_n,a)V_{n+1}(h_n,a,j)\} \geq \sup_{a\in\mathcal{A}}\{r_n(s_n,a) + \sum_{j\in\mathcal{S}} p_n(j|s_n,a)V_{n+1}^*(h_n,a,j)\} \geq \sup_{a\in\mathcal{A}}\{r_n(s_n,a) + \sum_{j\in\mathcal{S}} p_n(j|s_n,a)V_{n+1}^{\pi'}(h_n,a,j)\} \geq V_n^{\pi'}(h_n)$. Second prove for any $\varepsilon > 0$, there exists a $\pi \in \mathcal{D}^{\text{HD}}$ for which $V_n^{\pi'}(h_n) + (N-n)\varepsilon \geq V_n(h_n) \Rightarrow V_n^*(h_n) + (N-n)\varepsilon \geq V_n^{\pi'}(h_n) + (N-n)\varepsilon \geq V_n(h_n) \geq V_n^*(h_n)$. Construct a policy $\pi' = (\delta_1',\cdots,\delta_{N-1}')$ by choosing $\delta_n'(h_n)$ to satisfy $r_n(s_n,\delta_n'(h_n)) + \sum_{j\in\mathcal{S}} p_n(j|s_n,\delta_n'(h_n))V_{n+1}(h_n,\delta_n'(h_n)) + \varepsilon \geq V_n(h_n)$. By induction: $N: V_N^{\pi'}(h_N) = V_N(h_N)$. Assume that $V_t^{\pi'}(h_t) + (N-t)\varepsilon \geq V_t(h_t)$ for $t = n+1,\cdots,N$. For $t = n$, $V_n^{\pi'}(h_n) = r_n(s_n,\pi_n'(h_n)) + \sum_{j\in\mathcal{S}} p_n(j|s_n,\delta_n^{\pi'}(h_n))V_{n+1}^{\pi'}(h_n,\delta_n^{\pi'}(h_n),j) \geq V_n(h_n) - (N-n)\varepsilon$.) The equations yield that $\delta_t^*(h_t) \in \arg\max_{a\in\mathcal{A}}\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} p_t(s_t,a)V_{t+1}^*(h_t,a,j)\}$, which means it is HD, i.e. $U_N^*(s) = \sup_{\pi\in\mathcal{D}^{\text{HR}}} U_N^\pi(s) = \sup_{\pi\in\mathcal{D}^{\text{HD}}} U_N^\pi(s) \stackrel{?}{=} \sup_{\pi\in\mathcal{D}^{\text{MD}}} U_N^\pi(s)$.

- Let $V_t^*, t = 1,\cdots,N$ be solutions of Bellman Equations. Then (a) For each $t = 1,\cdots,N, V_t^*(h_t)$ depends on h_t only through s_t; (b) For any $\varepsilon > 0$, there exists an ε-optimal policy which is D and M; (c) Max can be achieved, it is optimal, which is MD. (Proof: (a): By induction, $V_N^*(h_N) = V_N^*(h_{N-1},a_{N-1},s) = r_N(s)$ for all $h_{N-1} \in \mathcal{H}_{N-1}$. Assume (a) is valid for $t = n+1,\cdots,N$. Then $V_n^*(h_n) = \sup_{a\in\mathcal{A}}\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} p_t(j|s_t,a)V_{t+1}^*(j)\} = V_n^*(s_t)$.)

- Backward Indcution (Dynamic Programming) Algorithm: 1. Set $t = N$ and $V_N^*(s_N) = r_N(s_N)$ for all $s_N \in \mathcal{S}$; 2. Substitute $t-1$ for t and compute $V_t^*(s_t)$ for each $s_t \in \mathcal{S}$: $V_t^*(s_t) = \max_{a\in\mathcal{A}}\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} p_t(j|s_t,a)V_{t+1}^*(s_t)\}$, set $\mathcal{A}_{s_t} = \arg\max_{a\in\mathcal{A}}\{r_t(s_t,a) + \sum_{j\in\mathcal{S}} p_t(j|s_t,a)V_{t+1}^*(s_t)\}$; 3. If $t = 1$, stop. Otherwise return to Step 2.

- Other remarks: (1) At time t, specialized $\mathcal{S}_t$ and $\mathcal{A}_s$, special structure for r_t and p_t; (2) $K = |\mathcal{S}|$ and $L = |\mathcal{A}|$, at eact t, only $(N-1)LK^2$ multiplications, ease computation and storage cost (because there are $(L^K)^{N-1}$ DM policies).

- Infinite-Horizon MDPs: Assumptions: Stationary reward and transition probabilities $r_t(s,a) \equiv r(s,a), p_t(j|s,a) \equiv p(j|s,a)$; Bounded rewards $|r(s,a)| \leq M < \infty$ for all $a \in \mathcal{A}$ and $s \in \mathcal{S}$; Discounting $\lambda, 0 \leq \lambda < 1$; Discrete state space $\mathcal{S}$. The expected total reward of policy $\pi = (\delta_1,\delta_2,\cdots) \in \mathcal{D}^{\text{HR}}: U^\pi(s) = \lim_{N\to+\infty} \mathbb{E}_s^\pi\{\sum_{t=1}^N \lambda^{t-1}r(X_t,Y_t)\} = \mathbb{E}_s^\pi\{\sum_{t=1}^{+\infty} \lambda^{t-1}r(X_t,Y_t)\}$. We say that a policy π^* is optimal when $U^{\pi^*}(s) \geq U^\pi(s)$ for each $s \in \mathcal{S}$ and all $\pi \in \mathcal{D}^{\text{HR}}$. Define the value of the MDP $U^*(s) = \sup_{\pi\in\mathcal{D}^{\text{HR}}} U^\pi(s)$. Let $U_\nu^\pi(s)$ denote the expected reward obtained by using π when the horizon ν is random. Then $U_\nu^\pi(s) = \mathbb{E}_s^\pi\{\mathbb{E}_{\nu\sim P}\sum_{t=1}^{\nu} r(X_t,Y_t)\}$. Let's recall geometric distribution with parameter $\lambda: \mathbb{P}(\nu = n) = (1-\lambda)\lambda^{n-1}, n = 1,2,\cdots$.

- Suppose ν has a GD(λ). Then $U^\pi(s) = U_\nu^\pi(s)$ for all $s \in \mathcal{S}$. (Proof: $\mathbb{E}_\nu^\pi(s) = \mathbb{E}_s^\pi\{\sum_{n=1}^{+\infty}\sum_{t=1}^{n} r(X_t,Y_t)(1-\lambda)\lambda^{n-1}\} = \mathbb{E}_s^\pi\{\sum_{t=1}^{+\infty}\sum_{n=t}^{+\infty} r(X_t,Y_t)(1-\lambda)\lambda^{n-1}\} = \mathbb{E}_s^\pi\{\sum_{t=1}^{+\infty} \lambda^{t-1}r(X_t,Y_t)\}$)

- Suppose $\pi \in \mathcal{D}^{\mathrm{HR}}$, then for each $s \in \mathcal{S}$, there exists a $\pi' \in \mathcal{D}^{\mathrm{MR}}$ for which $U^{\pi'}(s) = U^{\pi}(s)$. (Proof: Note that $U^{\pi}(s) = \mathbb{E}_s^{\pi}\{\sum\limits_{t=1}^{+\infty} \lambda^{t-1} r(X_t, Y_t)\} = \sum\limits_{t=1}^{+\infty}\sum\limits_{j\in\mathcal{S}}\sum\limits_{a\in\mathcal{A}} \lambda^{t-1} r(j,a) p^{\pi}(X_t = j, Y_t = a|X_1 = s)$. Fix $s \in \mathcal{S}$, so we only need to check $p^{\pi}(X_t = j, Y_t = a|X_1 = s) = p^{\pi'}(X_t = j, Y_t = a|X_1 = s)$. For each $j \in \mathcal{S}$ and $a \in \mathcal{A}$, define the randomized Markov decision rule δ_t' by $q_{\delta_t'(j)}(a) = p^{\pi}(Y_t = a|X_t = j, X_1 = s)$. Then $p^{\pi'}(Y_t = a|X_t = j) = p^{\pi}(Y_t = a|X_t = j, X_1 = s)$. Assume the conclusion holds for $t = 0, 1, \cdots, n-1$. Then $p^{\pi'}(X_n = j, Y_n = a|X_1 = s) = p^{\pi'}(Y_n = a|X_n = j, X_1 = s) p^{\pi'}(X_n = j|X_1 = s) = p^{\pi}(Y_n = a|X_n = j, X_1 = s) p^{\pi'}(X_n = j|X_1 = s)$. Then by induction assumption, $p^{\pi}(X_n = j|X_1 = s) = \sum\limits_{k\in\mathcal{S}}\sum\limits_{a\in\mathcal{A}} p^{\pi}(X_{n-1} = k, Y_{n-1} = a|X_1 = s) p(j|k,a) = \sum\limits_{k\in\mathcal{S}}\sum\limits_{a\in\mathcal{A}} p^{\pi'}(X_{n-1} = k, Y_{n-1} = a|X_1 = s) p(j|k,a) = p^{\pi'}(X_n = j|X_1 = s)$.)

- Vector express for MDP: δ MD, define $r_\delta(s)$ and $p_\delta(j|s)$ by $r_\delta(s) := r(s, \delta(s)), p_\delta(j|s) = p(j|s, \delta(s))$. Denote $r_\delta = (r_\delta(1), \cdots, r_\delta(|\mathcal{S}|))^T \in \mathbb{R}^{|\mathcal{S}|}, p_\delta = (p_\delta)_{(s,j)} = p(j|s, \delta(s))$. For MR δ, define $r_\delta(s) = \sum\limits_{a\in\mathcal{A}} q_{\delta(s)}(a) r(s,a), p_\delta(j|s) = \sum\limits_{a\in\mathcal{A}} q_{\delta(s)}(a) p(j|s,a)$. The (s,j)-th component of the t-step transition probability matrix p_π^t satisfies $p_\pi^t(j|s) = [p_{\delta_1} p_{\delta_2} \cdots p_{\delta_t}](j|s) = p^{\pi}(X_{t+1} = j|X_1 = s), \mathbb{E}_s^{\pi} g(X_t) = \sum\limits_{j\in\mathcal{S}} p_\pi^{t-1}(j|s) g(j) = (p_\pi^t g)_s$, and $U^{\pi} = \sum\limits_{t=1}^{+\infty} \lambda^{t-1} p_\pi^{t-1} r_{\delta_t} = r_{\delta_1} + \lambda p_{\delta_1}(r_{\delta_1} + \lambda p_{\delta_2} r_{\delta_2} + \cdots) = r_{\delta_1} + \lambda p_{\delta_1} U^{\pi_1}$. When π is stationary, $U = r_\delta + \lambda p_\delta U$.

- Define $\mathscr{L} U = \sup_{d\in\mathcal{D}^{\mathrm{MD}}}\{r_d + \lambda p_d U\}$. Suppose there exists a $U \in \mathcal{U}$ for which (a) $U \geq \mathscr{L} U$, then $U \geq U^*$; (b) $U \leq \mathscr{L} U$, then $U \leq U^*$; (c) $U = \mathscr{L} U$, then $U = U^*$. (Proof: (a) $U \geq \sup_{\delta\in\mathcal{D}^{\mathrm{MR}}}\{r_d + \lambda p_d U\} \geq r_{\delta_1} + \lambda p_{\delta_1} U \geq r_{\delta_1} + \lambda p_{\delta_1}(r_{\delta_2} + \lambda p_{\delta_2} U) \geq r_{\delta_1} + \lambda p_{\delta_1} r_{\delta_2} + \cdots + \lambda^{n-1} p_{\delta_1} p_{\delta_2} \cdots p_{\delta_{n-1}} r_{\delta_n} + \lambda^n p_\pi^n U \Rightarrow U - U^{\pi} \geq \lambda^n p_\pi^n U - \sum\limits_{k=n}^{+\infty} \lambda^k p_\pi^k r_{\delta_{k+1}} \geq 0$; (b) $U \leq \mathscr{L} U \Rightarrow U \leq r_d + \lambda p_d U + \varepsilon 1 \Rightarrow (I - \lambda p_d) U \leq r_d + \varepsilon 1 \Rightarrow U \leq (I - \lambda p_d)^{-1}(r_d + \varepsilon 1) = U^{\pi} + \varepsilon(1-\lambda)^{-1} 1_{|\mathcal{S}|}$.)

- If $0 \leq \lambda < 1$, $\mathscr{L}$ is a contraction mapping on $\mathcal{U}$. (Proof: Let u and v in $\mathcal{U}$. For each $s \in \mathcal{S}$, assume that $\mathscr{L} v(s) \geq \mathscr{L} u(s)$ and let $a_s^* = \arg\max_{a\in\mathcal{A}}\{r(s,a) + \sum\limits_{j\in\mathcal{S}} \lambda p(j|s,a) v(j)\}$. Then $0 \leq \mathscr{L} v(s) - \mathscr{L} u(s) \leq r(s, a_s^*) + \sum\limits_{j\in\mathcal{S}} \lambda p(j|s, a_j^*) v(j) - r(s, a_j^*) - \sum\limits_{j\in\mathcal{S}} \lambda p(j|s, a_s^*) u(j) = \lambda \sum\limits_{j\in\mathcal{S}} p(j|s, a_s^*)(v(j) - u(j)) \leq \lambda \sum\limits_{j\in\mathcal{S}} p(j|s, a_s^*) |\|u - v\| = \lambda \|u - v\|$.)

3 统计学习理论

- $(X, Y) \sim P \in \mathcal{P}$, definite $(X_1, Y_1), \cdots, (X_n, Y_n)$ i.i.d., $\mathcal{D}_n = \{(X_1, Y_1), \cdots, (X_n, Y_n)\}$, $\mathcal{R}_n(f) = \mathbb{E}_{(X,Y)\in\mathcal{D}_n} l(X, Y)$. An algorithm A is a mapping from $\mathcal{D}_n$ to function from $\mathcal{X} \to \mathcal{Y}$. Excess risk of A: $\mathcal{R}_P(A(\mathcal{D}_n)) - \mathcal{R}_P^*$. Expected error $\mathbb{E}[\mathcal{R}_P(A(\mathcal{D}_n))]$. An algorithm is called consistent in expectation for P iff $\mathbb{E}[\mathcal{R}_P(A(\mathcal{D}_n))] - \mathcal{R}_P^* \to 0$. PAC (probability approximately correct): for a given $\delta \in (0,1)$ and $\epsilon > 0$, $\mathbb{P}(\mathcal{R}_P(A(\mathcal{D}_n)) - \mathcal{R}_P^* \leq \epsilon) \geq 1 - \delta$.

- 回归: $g(x) = \mathbb{E}[Y|X = x], g_n(x, \mathcal{D}_n) = g_n(x), \mathbb{E}\{|g_n(X) - Y|^2|\mathcal{D}_n\} = \int_{\mathbb{R}^d} |g_n(x) - g(x)|^2 \mu(\mathrm{d}x) + \mathbb{E}|g(X) - Y|^2$. A sequence of regression function estimates $\{g_n\}$ is called weakly consistent for a certain distribution of (X, Y) if $\lim\limits_{n\to+\infty} \mathbb{E}\{\int [g_n(x) - g(x)] \mu(\mathrm{d}x)\} = 0$; strongly consistent for a certain distribution if $\lim\limits_{n\to+\infty} \int [g_n(x) - g(x)]^2 \mu(\mathrm{d}x) = 0$ with probability 1; weakly universally consistent if for all distributions of (X, Y) with $\mathbb{E}[Y^2] < \infty, \cdots$; strongly universally consistent $\cdots$.

- Penalized model: $g_n = \arg\min_f\{\frac{1}{n}\sum\limits_{i=1}^n |f(X_i) - Y_i|^2 + J_n(f)\}$. Penalized term for f: $J_n(f) = \lambda_n \int |f''(t)|^2 \mathrm{d}t, J_{n,k}(f) = \lambda_n \int \sum\limits_{t_1,\cdots,t_k\in\{1,\cdots,d\}} |\frac{\partial f^k}{\partial x_{t_1} \cdots \partial x_{t_d}}|^2 \mathrm{d}t$.

- Curse of dimensionality: let $X, X_1, \cdots, X_n$ i.i.d. $\mathbb{R}^d$ uniformly distributed in $[0,1]^d$. $d_\infty(d,n) = \mathbb{E}\{\min_{i=1,\cdots,n} \|X - X_i\|_\infty\} = \int_0^\infty \mathbb{P}\{\min_{i=1,\cdots,n} \|X - X_i\|_\infty > t\} \mathrm{d}t = \int_0^\infty (1 - \mathbb{P}\{\min_{i=1,\cdots,n} \|X - X_i\|_\infty < t\}) \mathrm{d}t$. Since $\mathbb{P}\{\min_i \|X - X_i\|_\infty < t\} \leq n\mathbb{P}(\|X - X_1\|_\infty \leq t) \leq n(2t)^d$, 原式 $\geq \frac{d}{2(d+1)} n^{-\frac{1}{d}}$.

- No-Free lunch: Let $\{a_n\}$ be a sequence of positive numbers converging to 0. For every sequence of regression estimates, there exists a distribution of (X, Y) such that X is uniformly distributed on $[0,1]$, $Y = g(X)$, g is ± 1 valued, and $\limsup_{n\to+\infty} \frac{\mathbb{E}\|g_n - g\|^2}{a_n} \geq 1$. (Proof: Let $\{p_i\}$ be a probability distribution and let $\mathscr{A} = \{\mathscr{A}_j\}$

be a partition of $[0,1]$ such that $\mathscr{A}_j$ is an interval of length p_j. Consider regression function indexed by a parameter c, $c=(c_1,c_2,\cdots)$ where $c_j\in\{\pm1\}$. Define $g^{(c)}:[0,1]\to\{-1,1\}$ by $g^{(c)}(x)=c_j$ if $x\in\mathscr{A}_j$ and $Y=g^{(c)}(x)$. For $x\in\mathscr{A}_j$, define $\bar{g}_n(x)=\frac{1}{p_j}\int_{\mathscr{A}_j}g_n(z)\mu(\mathrm{d}z)$ to be the projection of g_n on $\mathscr{A}$. Then $\int_{\mathscr{A}_j}|g_n(x)-g^{(c)}(x)|^2\mu(\mathrm{d}x)=\int_{\mathscr{A}_j}|g_n(x)-\bar{g}_n(x)|^2\mu(\mathrm{d}x)+\int_{\mathscr{A}}|\bar{g}_n(x)-g^{(c)}(x)|^2\mu(\mathrm{d}x)\geq\int_{\mathscr{A}}|\bar{g}_n(x)-g^{(c)}(x)|^2\mu(\mathrm{d}x)$. Set $\hat{c}_{nj}=1$ if $\int_{\mathscr{A}_j}g_n(z)\mu(\mathrm{d}z)\geq0$; $=-1$, otherwise. For $x\in\mathscr{A}_j$, if $\hat{c}_{nj}=1$ and $c_j=-1$, then $\bar{g}_n(x)\geq0$ and $g^{(c)}(x)=-1$, implying $|\bar{g}_n(x)-g^{(c)}(x)|\geq1$; if $\hat{c}_{nj}=-1$ and $c_j=1$, then $\bar{g}_n(x)<0$ and $g^{(c)}(x)=1\Rightarrow|\bar{g}_n(x)-g^{(c)}(x)|^2\geq1$. Therefore $\int_{\mathscr{A}}|\bar{g}_n(x)-g^{(c)}(x)|^2\mu(\mathrm{d}x)\geq1_{\{\hat{c}_{nj}\neq c_j\}}\int_{\mathscr{A}_j}1\mu(\mathrm{d}x)\geq1_{\{\hat{c}_{nj}\neq c_j\}}p_j\geq1_{\{\hat{c}_{nj}\neq c_j\}}1_{\{\mu_n(\mathscr{A}_j)=0\}}p_j\Rightarrow\mathbb{E}\{\int|g_n(x)-g^{(c)}(x)|^2\mu(\mathrm{d}x)\}\geq\sum_{j=1}^{+\infty}\mathbb{P}(\hat{c}_{nj}\neq c_j,\mu_n(\mathscr{A}_j)=0)p_j:=R_n(c)$. Now we randomize c. Let $C_1,C_2,\cdots$ be a sequence of i.i.d. random variables independent of $X_1,X_2,\cdots$ which satisfy $\mathbb{P}(c_1=1)=\mathbb{P}(c_1=-1)=\frac{1}{2}$. Thus $\mathbb{E}R_n(C)=\sum_{j=1}^{+\infty}\mathbb{E}\mathbb{P}(\hat{C}_{nj}\neq C_j,\mu_n(\mathscr{A}_j)=0)p_j\overset{\text{重期望}}{=}\sum_{j=1}^{+\infty}\mathbb{E}\{1_{\{\mu_n(\mathscr{A}_j)=0\}}\mathbb{P}(\hat{C}_{nj}\neq C_j|X_1,\cdots,X_n)\}p_j=\frac{1}{2}\sum_{j=1}^{+\infty}\mathbb{P}(\mu_n(\mathscr{A}_j)=0)p_j=\frac{1}{2}\sum_{j=1}^{+\infty}(1-p_j)^np_j$. On the other hand, $R_n(c)\leq\sum_{j=1}^{+\infty}\mathbb{P}(\mu_n(\mathscr{A}_j)=0)p_j=\sum_{j=1}^{+\infty}(1-p_j)^np_j\Rightarrow\frac{R_n(c)}{\mathbb{E}R_n(C)}\leq2$. By Fatou's lemma, $\mathbb{E}\{\limsup_{n\to+\infty}\frac{R_n(C)}{\mathbb{E}R_n(C)}\}\geq\limsup_{n\to+\infty}\{\frac{R_n(C)}{\mathbb{E}R_n(C)}\}=1$, which implies that there exists $c\in C$ such that $\limsup_{n\to+\infty}\frac{R_n(C)}{\mathbb{E}R_n(C)}\geq1\Rightarrow\limsup_{n\to+\infty}\frac{\mathbb{E}\{\int|g_n(x)-g(x)|^2\mu(\mathrm{d}x)\}}{\frac{1}{2}\sum_{j=1}^{+\infty}(1-p_j)^np_j}\geq1$. Let $\{a_n\}$ be a sequence of positive numbers converging to 0 with $\frac{1}{2}\geq a_1\geq a_2\geq\cdots$, then there exists a probability $\{p_j\}$ such that $\sum_{j=1}^{+\infty}(1-p_j)^np_j\geq a_n,\forall n$.)

- Minimax lower Bounds: (a) The sequence of positive numbers a_n is called the lower minimax rate of convergence for the $\mathcal{P}$ if $\liminf_{n\to+\infty}\inf_{g_n}\sup_{P\in\mathcal{P}}\frac{\mathbb{E}\{\|g_n-g\|^2\}}{a_n}=c_1>0$. (b) a_n is called optimal rate of convergence for the class $\mathcal{P}$ if it is a lower minimax rate of convergence and there is an estimate g_n such that $\limsup_{n\to+\infty}\sup_{P\in\mathcal{P}}\frac{\mathbb{E}\|g_n-g\|^2}{a_n}=c_n<\infty$.

- Smoothness: Let $q=k+\beta$ for some $k\in\mathbb{N}$ and $0<\beta\leq1$ and let $\rho>0$. A function $f:\mathbb{R}^d\to\mathbb{R}$ is called (q,ρ)-smooth if for every $\alpha=(\alpha_1,\cdots,\alpha_d),\alpha_i\in\mathbb{N},\sum_{i=1}^{d}\alpha_i=k$, the partial derivative $\frac{\partial^kf}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}$ exists and satisfies $\left|\frac{\partial^kf}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}(x)-\frac{\partial^kf}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}(z)\right|\leq\rho\|x-z\|^\beta$. Let $\mathscr{F}^{(q,\rho)}$ be the set of all (q,ρ)-smooth functions f. Let $\mathcal{P}^{(q,\rho)}$ be the class of distributions (X,Y) such that (i) X is uniformly distributed on $[0,1]^d$; (ii) $Y=g(X)+N$, where $X\perp\!\!\!\perp N$, and N is standard normal; (iii) $g\in\mathscr{F}^{q,\rho}$.

- Let u be an l-dimensional real vector, let C be a zero means random variables takeing values in $\{-1,1\}$ and let N be an l-dimensional standard normal independent of C. Set $Z=Cu+N$. Then the error probability of the Bayesian decision for C based on Z is $\mathcal{R}^*=\min_{g:\mathbb{R}^l\to\mathbb{R}}\mathbb{P}(g(Z)\neq C)=\Phi(-\|u\|)$. (Proof: $\mathbb{P}(C=1)=\mathbb{P}(C=-1)=\frac{1}{2},\mathbb{P}(Z|C=1)=\mathcal{N}(u,I),\mathbb{P}(Z|C=-1)=\mathcal{N}(-u,I)$. By the Bayes formula, $\mathbb{P}(C=1|Z=z)=\frac{\mathbb{P}(C=1)\mathbb{P}(Z|C=1)}{\mathbb{P}(C=1)\mathbb{P}(Z|C=1)+\mathbb{P}(C=-1)\mathbb{P}(Z|C=-1)}=\frac{1}{1+\exp(\frac{\|Z-u\|^2}{2}-\frac{\|Z+u\|^2}{2})}=\frac{1}{1+\exp(-2Z^Tu)}$. Therefore, the optimal Bayes decision is $g^*(Z)=\mathrm{sgn}(Z^Tu)$, the risk $\mathcal{R}^*=\mathbb{P}(g^*(Z)\neq C)=\mathbb{P}(Z^Tu<0,C=1)+\mathbb{P}(Z^Tu>0,C=-1)=\mathbb{P}(\|u\|^2+u^TN<0,C=1)+\mathbb{P}(-\|u\|^2+u^TN>0,C=-1)=\frac{1}{2}\mathbb{P}(u^TN\leq-\|u\|^2)+\frac{1}{2}\mathbb{P}(u^TN>\|u\|^2)=\Phi(-\|u\|)$.)

- For the class $\mathcal{P}^{(q,\rho)}$, the sequence $a_n=n^{-\frac{2q}{2q+d}}$ is a lower minimax rate of convergence. In particular,

$$\liminf_{n\to\infty}\inf_{g_n}\sup_{P_{(X,Y)}\in\mathcal{P}^{(q,\rho)}}\frac{\mathbb{E}\|g_n-g\|^2}{\rho^{\frac{2d}{2q+d}}n^{-\frac{2q}{2q+d}}}\geq c_1>0.$$

证明分为 4 步. Step 1: 构造一个辅助函数 $g^{(c)}$. Set $M_n=\lceil(\rho^2n)^{\frac{1}{2q+d}}\rceil$. Partition $[0,1]^d$ by M_n^d cubes $\{A_{n,j}\}$ of side length $\frac{1}{M_n}$ and with centers $\{a_{n,j}\}$. Choose a function $\bar{f}:\mathbb{R}^d\to\mathbb{R}$ such that the support of $\bar{f}$ is a subset of $[-\frac{1}{2},\frac{1}{2}]^d,\int\bar{f}^2(x)\mathrm{d}x>0$ and $\bar{f}\in\mathscr{F}^{(q,2^{\beta-1})}$. Define $f:\mathbb{R}^d\to\mathbb{R}$ by $f=\rho\bar{f}$. Let $c_n=(c_{n,1},\cdots,c_{n,M_n^d})\in\mathcal{C}_n$ take values in $\{\pm1\}$. $g^{(c_n)}(x)=\sum_{j=1}^{M_n^d}c_{n,j}f_{n,j}(x)$ where $f_{nj}(x)=M_n^{-q}f(M_n(x-a_{n,j}))$.

Step 2: 证明 $g^{(c_n)}\in\mathscr{F}^{(q,\rho)}$. Let $\alpha=(\alpha_1,\cdots,\alpha_d),\alpha_i\in\mathbb{N}$ and $\sum_{j=1}^{d}\alpha_j=k$. Set $D^\alpha=\frac{\partial^k}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}$. If $x,z\in A_{n,j}$, $|D^\alpha g^{c_n}(x)-D^\alpha g^{(c_n)}(z)|=|c_{n,k}||D^\alpha f_{n,j}(x)-D^\alpha f_{n,j}(z)|\leq\rho\|x-z\|^\beta$. If $x\in A_{n,i},z\in A_{n,j}$, choose $\bar{x},\bar{z}$ on the line between x and z such that $\bar{x}$ is on the boundary of $A_{n,i}$ and $\bar{z}$ is on the boundary of $A_{n,j}$. $|D^\alpha g^{(c_n)}(x) - D^\alpha g^{(c_n)}(z)| \leq |c_{n,i} D^\alpha f_{n,i}(x)| + |c_{n,j} D^\alpha f_{n,j}(z)| = |c_{n,i}||D^\alpha f_{n,i}(x) - D^\alpha f_{n,i}(\bar{x})| + |c_{n,j}||D^\alpha f_{n,j}(z) - D^\alpha f_{n,j}(\bar{z})| \leq \rho 2^{\beta-1}(\|x-\bar{x}\|^\beta + \|z-\bar{z}\|^\beta) = \rho 2^\beta(\frac{\|x-\bar{x}\|^\beta}{2} + \frac{\|z-\bar{z}\|^\beta}{2}) \leq \rho 2^\beta(\frac{\|x-\bar{x}\|}{2} + \frac{\|z-\bar{z}\|}{2})^\beta \leq \rho\|x-z\|^\beta$.

Step 3: Prove that $\liminf\limits_{n\to+\infty} \inf_{g_n} \sup_{Y=g^{(c)}(X)+N, c\in\mathcal{C}_n} \frac{M_n^{2q}}{\rho^2}\mathbb{E}\|g_n - g^{(c)}\|^2 > 0$. $\{f_{n,j}\}$ forms a set of orthogonal basis. Let g_n be an arbitrary estimate, and the projection $\bar{g}_n$ of g_n to $\{g^{(c)} : c \in \mathcal{C}_n\}$ is given by $\bar{g}_n = \sum\limits_{j=1}^{M_n} \tilde{c}_{n,j} f_{n,j}(x)$. $\|g_n - g^{(c)}\|^2 = \|g_n - \bar{g}_n\|^2 + \|g_n - g^{(c)}\|^2 \geq \|\bar{g}_n - g^{(c)}\|^2 = \sum\limits_{j=1}^{M_n^d} \int_{A_{n,j}} (\tilde{c}_{n,j} f_{n,j}(x) - c_{n,j} f_{n,j}(x))^2 \mathrm{d}x = \sum\limits_{j=1}^{M_n^d} \int_{A_{n,j}} (\tilde{c}_{n,j} - c_{n,k})^2 f_{n,j}^2(x)\mathrm{d}x = \int f^2(x)\mathrm{d}x \sum\limits_{j=1}^{M_n^d} (\tilde{c}_{n,j} - c_{n,j})^2 \frac{1}{M_n^{2q+d}}$. Define $\bar{c}_{n,j} = \mathrm{sgn}(\tilde{c}_{n,j})$, $|\tilde{c}_{n,j} - c_{n,j}| \geq \frac{|\bar{c}_{n,j} - c_{n,j}|}{2} \Rightarrow \|g_n - g^{(c)}\|^2 \geq \int f^2(x)\mathrm{d}x \frac{1}{4} \frac{1}{M_n^{2q+d}} \sum\limits_{j=1}^{M_n^d} (\bar{c}_{n,j} - c_{n,j})^2 = \frac{\rho^2}{M_n^{2q}} \int \bar{f}^2(x)\mathrm{d}x \frac{1}{M_n^d} \sum\limits_{j=1}^{M_n^d} 1_{\{\bar{c}_{n,j} \neq c_{n,j}\}}$.

Step 4: Prove that $\liminf\limits_{n\to+\infty} \inf_{\bar{c}_n} \sup_{c_n} \frac{1}{M_n^d} \sum\limits_{j=1}^{M_n^d} \mathbb{P}(\bar{c}_{n,j} \neq c_{n,j}) > 0$. Now we randomize c_n. Let $c_{n,1}, \cdots, c_{n,M_n^d}$ be i.i.d. random variables independent of $(X_1, N_1), \cdots, (X_n, N_n)$, $\mathbb{P}(C_{n,1} = 1) = \mathbb{P}(C_{n,1} = -1) = \frac{1}{2}$. $\bar{c}_{n,j}$ can be interpreted as a decision on $C_{n,j}$ using $\mathcal{D}_n$. Let $\bar{C}_{n,j} = 1$ if $\mathbb{P}(\bar{C}_{n,j} = 1|\mathcal{D}_n) \geq \frac{1}{2}$. Therefore, $\inf_{\bar{c}_n} \sup_{c_n} \frac{1}{M_n^d} \sum\limits_{j=1}^{M_n^d} \mathbb{P}(\bar{c}_{n,j} \neq c_{n,,j}) \geq \inf_{\bar{c}_n} \frac{1}{M_n^d} \sum\limits_{j=1}^{M_n^d} \mathbb{P}(\bar{c}_{n,j} \neq C_{n,j}) \geq \frac{1}{M_n^d} \sum\limits_{j=1}^{M_n^d} \mathbb{P}(\bar{C}_{n,j} \neq C_{n,j}) = \mathbb{P}(\bar{C}_{n,1} \neq C_{n,1}) = \mathbb{E}\{\mathbb{P}(\bar{C}_{n,1} \neq C_{n,1}|X_1, \cdots, X_n)\}$. Let $X_{i_1}, \cdots, X_{i_t}$ be those $X_i \in A_{n,1}$, $(Y_{i,1}, \cdots, Y_{i_t}) = C_{n,1}(f_{n,1}(X_{i_1}), \cdots, f_{n,1}(X_{i_t})) + (N_{i_1}, \cdots, N_{i_t})$. By the latest lemma, $\mathbb{E}\{\mathbb{P}(\bar{C}_{n,1} \neq C_{n,1}|X_1, \cdots, X_n)\} = \mathbb{E}\Phi\left(-\sqrt{\sum\limits_{r=1}^{t} f_{n,1}^2(X_{i_r})}\right) = \mathbb{E}\Phi\left(-\sqrt{\sum\limits_{i=1}^{n} f_{n,1}^2(X_i)}\right) \geq \Phi\left(-\sqrt{\mathbb{E}\sum\limits_{i=1}^{n} f_{n,1}^2(X_i)}\right) \geq \Phi(-\sqrt{\int f^2(x)\mathrm{d}x}) > 0$.