Asymptotic Properties for Bayesian Neural Network in Besov Space

김성민

February 16, 2023

1 레마1

정의 1 (Hellinger distance)**.**

$$\rho(P,Q) = \int \sqrt{dPdQ}, \quad h^2(P,Q) := \frac{1}{2}\int (\sqrt{dP} - \sqrt{dQ})^2 = 1 - \rho(P,Q)$$

여기서 P_c 와 Q_c 를 중심으로 하는 공은 각각 다음과 같이 정의할 수 있다.

$$A = \{P : \rho(P,P_c) \geq \alpha\},$$
$$B = \{Q : \rho(Q,Q_c) \geq \beta\}.$$

레마 1 (Lemma8 **?**)**.** *Asumme that the Hellinger balls A and B are disjoint. Then there is a uniquely determined pair $P_m \in A$ and $Q_m \in B$ such that*

$$\rho_m = \rho(P_m, Q_m) = \rho(A,B) := \sup\{\rho(P,Q) : P \in A,\ Q \in B\}.$$

The measures P_m and Q_m are mutually absolute continuous and equivalent to $P_c + Q_c$.

참고 1 (Most unfavorable set)**.** *Let $\rho_c = \rho(P_c, Q_c) = \cos w$, $\alpha = \cos\epsilon$ and $\beta = \cos\eta$. Then most unfavorable set for the sets (A,B) are as following:*

$$\sqrt{dP_m} = (\sin w)^{-1}[\sin(w-\epsilon)\sqrt{dP_c} + \sin\epsilon\sqrt{dQ_c}]$$
$$\sqrt{dQ_m} = (\sin w)^{-1}[\sin(w-\eta)\sqrt{dQ_c} + \sin\eta\sqrt{dP_c}].$$

정리 1 (Theorem1 **?**). *With the notation and assumptions of Lemma1, let ϕ be a measurable funciton defined on $(\mathcal{X}, \mathcal{A})$ with the following properties:*

1. *For the measure $P_c + Q_c$, the function ϕ is equivalent to $\sqrt{dQ_m/dP_m}$.*

2. *Let $m = essinf\phi$ and $M = esssup\phi$ for $P_c + Q_c$. Then $m \le \phi(x) \le M$ for all $x \in \mathcal{X}$.*

For such a ϕ, one has $\int \phi dP \le \rho_m$ and $\int \phi^{-1} dQ \le \rho_m$ for all $P \in A$ and $Q \in B$.

Consider two probability measures P_c and Q_c. Let w be the angle defined by $w = \arccos \rho(P_c, Q_c)$ and let ξ be a number $0 < \xi \le \frac{1}{3}$. Let A be the ball

$$A = \{P : \rho(P, P_c) \ge \cos \xi w\} = \Big\{P : h^2(P, P_c) \le 2 \sin^2 \frac{\xi w}{2}\Big\}.$$

Let B denote the ball $B = \{Q : h^2(Q, Q_c) \le 2 \sin^2(\xi w/2)\}$. With this notation the function ϕ can be written

$$\phi = \frac{[\sin(1-\xi)w]\sqrt{dQ_c} + [\sin \xi w]\sqrt{dP_c}}{[\sin \xi w]\sqrt{dQ_c} + [\sin(1-\xi)w]\sqrt{dP_c}}.$$

The least favorable pair (P_m, Q_m) is given by $\sqrt{dP_m} = (\sin w)^{-1}\{[\sin \xi w]\sqrt{dQ_c} + [\sin(1-\xi)w]\sqrt{dP_c}\}$ and by the analogous expression for $\sqrt{dQ_m}$.

레마 2 (Lemma 12 **?**). *Let $P_c, Q_c, \xi,$ and w be as described, with $\xi \le \frac{1}{3}$. Let $h = h(P_c, Q_c)$, $d = h(P_m, Q_m)$, $r = h(P_m, Q_m) = \sqrt{2}\sin(\xi w/2)$, and $R(\xi) = [\sin(1-2\xi)\pi/4][(1-2\xi)\sin \pi/4]^{-1}$. Let P be an arbitrary probability measure. Let $C = h(P, P_c)$ and $z = r^2/C^2$. Let y be the solution of the equation*

$$y^2(y+1)R^2(\xi) = \frac{4}{z}.$$

Then

$$\int \phi dP \le (1 - d^2) + 2\Big(\frac{1-2\xi}{\xi}\Big)\Big[1 + \frac{3}{y} + \frac{1}{y(y+1)}\Big][1-z]^+ C^2 \tag{1}$$

$$\le 1 - (1-2\xi)^2 h^2 + 4\Big(\frac{1-2\xi}{\xi}\Big)C^2. \tag{2}$$

Proof.

- $P \in A = \{P : h(P, P_c) \le r\}$ 인 경우를 생각하자. 이 부분은 정리1에 의해 성립된다.

$$\int \phi dP \le \rho_m = 1 - d^2 \le 1 - h^2.$$

	$\phi > 1$	$\phi < 1$
$\sqrt{f} \geq x$	S	E
$1 \leq \sqrt{f} < x$	T'	F
$0 \leq \sqrt{f} < 1$	S'	T''

따라서 $C = h(P, P_C) > r$ 를 가정할 수 있다.

$P^* := (1-\beta)P + \beta P_c$ with $\beta = C^{-2}[C^2 - r^2]$ 이라고 정의하면 다음을 얻을 수 있다.

$$\begin{aligned} h^2(P^*, P) &= 1 - \int \sqrt{dP^* dP_c} \\ &= 1 - \int \sqrt{(1-\beta)dPdP_c + \beta dP_c dP_c} \\ &\leq 1 - \int (1-\beta)\sqrt{dPdP_c} - \beta\sqrt{dP_c dP_c} \\ &= (1-\beta)C^2 = r^2 \end{aligned}$$

이므로 $P^* \in A$ 임을 알 수 있다. 따라서 $\int \phi d(P - P^*) = \beta \int \phi d(P - P_c)$ 의 상한을 구하면 충분하다.

- 적절한 함수 f가 있어서 $dP = fdP_c + dP'$로 쓸 수 있다. ($f = 0$ on Ω_c^c, $\quad \Omega_c : P_c$의 토대) 여기서 P'은 P의 일부로 P_c-singular 인 파트이다. 또한 $x > 1$를 가정하자. 이 때 아래 표처럼 집합을 나눌 수 있다.
  - E, F 는 $\int (\phi - 1)d(P - P_c) \leq 0$ 임을 쉽게 보일 수 있다.
  - S 상에서 $f - 1 \leq \frac{x+1}{x-1}(\sqrt{f} - 1)^2$ 이다.
  $$\begin{aligned} \int_S (\phi - 1)d(P - P_c) &\leq \frac{x+1}{x-1}\frac{1-2\xi}{\xi}\int_S (\sqrt{f} - 1)^2 dP_c + \frac{1-2\xi}{\xi}\int_S dP' \\ &= \frac{x+1}{x-1}\frac{1-2\xi}{\xi}\int_S (\sqrt{f} - 1)^2 dP_c \end{aligned}$$
  - $$\begin{aligned} \int_{S'} (\phi - 1)d(P - P_c) &= \int_{S' \cap \Omega_c} (\phi - 1)d(P - P_c) + \int_{S' \cap \Omega_c^c} (\phi - 1)d(P - P_c) \\ &\leq \int_{S' \cap \Omega_c} (\phi - 1)d(P - P_c) \\ &\leq \frac{1-2\xi}{\xi}\int_{S'} (\sqrt{dP} - \sqrt{dP_c})^2 \quad (\because\ \Omega_c \text{에서 } d(P - P_c) = (\sqrt{dP} - \sqrt{dP_c})^2.) \end{aligned}$$
  - 이제 T' 집합을 고려하자. 이 집합에서 다음이 성립한다.

    $\phi > 1 \iff \sqrt{dQ_c} > \sqrt{dP_c}$

    $\sin(1-\xi)w + \sin \xi w \geq \sin w.$

  $$\implies \quad \sqrt{dP_m} = (\sin w)^{-1}\{[\sin \xi w]\sqrt{dQ_c} + [\sin(1-\xi)w]\sqrt{dP_c}\} > \sqrt{dP_c}\frac{\sin(1-\xi)w + \sin \xi}{\sin w} \geq \sqrt{dP_c}$$

이를 이용해서 다음과 같은 부등식을 얻을 수 있다.

$$\begin{aligned}\int_{T'}(\phi-1)d(P-P_c) &= \int_{T'}(\phi-1)(f-1)dP_c \\ &= \int_{T'}(\phi-1)(\sqrt{f}+1)(\sqrt{f}-1)dP_c \\ &\leq (x+1)\int_{T'}(\phi-1)(\sqrt{f}-1)\sqrt{dP_c dP_m}.\end{aligned}$$

– T'' 집합을 고려하자. 이 집합에서는 다음이 성립한다.

$$\sqrt{dP_c} \leq \frac{\sin w}{\sin(1-\xi)w}\sqrt{dP_m} \leq \frac{1}{1-\xi}\sqrt{dP_m}$$

$$\sqrt{f}+1 \leq 2.$$

이를 이용해서 다음과 같은 부등식을 얻을 수 있다.

$$\begin{aligned}\int_{T''}(\phi-1)d(P-P_c) &= \int_{T''}(\phi-1)(f-1)dP_c + \int_{T''}(\phi-1)dP' \\ &\leq \int_{T''}(\phi-1)(f-1)dP_c \\ &\leq \frac{2}{1-\xi}\int_{T''}(\phi-1)(\sqrt{f}-1)\sqrt{dP_c dP_m}.\end{aligned}$$

– $T = T' \cup T''$ 이고 $\lambda = \max(x+1, \dfrac{2}{1-\xi})$ 라고 하면 다음을 얻을 수 있다.

$$\begin{aligned}\int_T(\phi-1)d(P-P_c) &\leq \lambda\int_T(\phi-1)(\sqrt{f}-1)\sqrt{dP_m dP_c} \\ &\leq \lambda\Big\{\int_T(\phi-1)^2dP_m\int_T(\sqrt{f}-1)^2dP_c\Big\}^{\frac{1}{2}} \quad \text{(코시-슈바르츠 부등식)}.\end{aligned}$$

– 다음과 같은 사실을 알 수 있다.

$$\begin{aligned}\int_X(\sqrt{f}-1)^2dP_c &= \int_X(\sqrt{fdP_c}-\sqrt{dP_c})^2 \\ &= \int_{X\cap\Omega_c}(\sqrt{fdP_c}-\sqrt{dP_c})^2 + \int_{X\cap\Omega_c^c}(\sqrt{fdP_c}-\sqrt{dP_c})^2 \\ &= \int_{X\cap\Omega_c}(\sqrt{fdP_c}-\sqrt{dP_c})^2 \\ &\leq \int_X(\sqrt{dP}-\sqrt{dP_c})^2 \qquad (3)\end{aligned}$$

– $I(X) := \int_X(\phi-1)^2dP_m$ 이고 $J(X) := \int_X(\sqrt{f}-1)^2dP_c$, $A = \dfrac{x+1}{x-1}\dfrac{1-2\xi}{\xi}$ 라고 하면 다음과 같은 부등식을 얻을 수 있다.

$$\begin{aligned}\int(\phi-1)d(P-P_c) &\leq A\int_S(\sqrt{f}-1)^2dP_c + A\int_{S'}(\sqrt{dP}-\sqrt{dP_c})^2 + \lambda(I(T)J(T))^{\frac{1}{2}} \\ &\leq A[\int(\sqrt{dP}-\sqrt{dP_c})^2 - J(T)] + \lambda(I(T)J(T))^{\frac{1}{2}} \quad (by\ (3)) \\ &= A(2C^2 - J(T)) + \lambda(I(T)J(T))^{\frac{1}{2}}\end{aligned}$$

- 여기서 $\xi \le \dfrac{1}{3}$ 이면 $x+1 \ge \dfrac{2}{1-\xi}$ 인 적절한 x를 선택할 수 있고, 따라서 $\lambda = x+1$ 이다.

  또한 간단한 계산을 통해서 $J(T) = \dfrac{\lambda^2 I(T)}{4A^2}$ 일 때 우항이 최대가 됨을 알 수 있다.

$$\int (\phi - 1) d(P - P_c) \le 2AC^2 + \frac{\lambda^2 I}{4A} = 2\Big(\frac{1-2\xi}{\xi}\Big)\frac{x+1}{x-1}C^2 + \frac{\xi}{4(1-2\xi)}(x^2-1)I(T).$$

- $R(\xi, w) = \dfrac{\sin(1-2\xi)\frac{w}{2}}{(1-2\xi)\sin\frac{w}{2}}$ 이라고 정의하면 $d^2 = R^2(\xi, w)(1-2\xi)^2 h^2$ 이 된다.

  또한 $R(\xi, w)$ 은 w 에 대한 증가함수이므로,

$$R(\xi, w) \le R(\xi) = \frac{\sin(1-2\xi)\frac{\pi}{4}}{(1-2\xi)\sin\frac{\pi}{4}}$$

  가 성립한다.

  따라서, $\displaystyle\int (\phi - 1) d(P - P_c) \le \frac{1-2\xi}{\xi}\Big\{2\frac{x+1}{x-1}C^2 + \frac{x^2-1}{2}R^2(\xi)\xi^2 h^2\Big\}$ 이다.

  $(\xi^2 h^2 = 2\xi^2 \sin^2 \dfrac{w}{2} \le 2\sin^2 \dfrac{\xi w}{2} = r^2)$

  여기서 $z = \dfrac{r^2}{C^2}$ 이고 $\beta = C^{-2}(C^2 - r^2) = 1 - z$ 라고 놓으면

$$\beta \int (\phi - 1) d(P - P_c) \le C^2 \frac{1-2\xi}{\xi}\Big\{2(1 + \frac{2}{y}) + \frac{y(y+2)}{2}R^2(\xi) z\Big\}(1-z)$$

  where $y = x - 1$ should be such that $y \ge \dfrac{2\xi}{1-\xi}$.

-
  - 우항을 최소화하는 y는 다음의 해이다.

$$y^2(y+1) = \frac{4}{R^2(\xi) z}$$

  - $P \notin A$ 조건 때문에 $z < 1$ 이다.
  - $1 \le R^2(\xi) \le \dfrac{\pi^2}{8}$

  따라서 최소화하는 y 는 만족해야하는 $x + 1 = y + 2 \ge \dfrac{2}{2-\xi}$ 를 $\xi \le \dfrac{1}{3}$ 이기만 하면 만족한다.

-

$$\begin{aligned}
\beta \int (\phi - 1) d(P - P_c) &\Big(\le C^2 \frac{1-2\xi}{\xi}\Big\{2(1 + \frac{2}{y}) + \frac{y(y+2)}{2}R^2(\xi) z\Big\}(1-z)\Big) \\
&\le C^2 \frac{1-2\xi}{\xi}\Big\{2(1 + \frac{2}{y}) + \frac{y(y+2)}{2}\frac{4}{y^2(y+1)}\Big\}(1-z) \\
&= 2C^2\Big(\frac{1-2\xi}{\xi}\Big)Q(y) \\
&\le 4C^2\Big(\frac{1-2\xi}{\xi}\Big). \qquad (4)
\end{aligned}$$

부등식 (4) 은 다음에 의해 성립한다.

$$\begin{aligned}Q(y) &= (1-z)\Big[1+\frac{3}{y}+\frac{1}{y(y+1)}\Big]\\ &= \Big[1+\frac{3}{y}+\frac{1}{y(y+1)}\Big]\Big[1-\frac{4}{y^2(y+1)R^2(\xi)}\Big]\\ &\le 2\end{aligned}$$

$R^2(\xi) \le \dfrac{\pi^2}{8}$ 이기 때문에 Q를 최대화하는 y는 $y^2(y+1) \ge 4$를 만족한다. 실제로 계산하면 $y \in (2.2, 2.3)$ 에서 $Q(y) \le 2$ 이다.

- 따라서 다음과 같이 결론을 지을 수 있다.

$$\begin{aligned}\int \phi dP &\le \int \phi dP^* + \beta \int \phi d(P-P_c)\\ &\le \rho_m + 4C^2\Big(\frac{1-2\xi}{\xi}\Big)\\ &= 1-d^2+4C^2\Big(\frac{1-2\xi}{\xi}\Big)\\ &\le 1-(1-2\xi)^2h^2+4\Big(\frac{1-2\xi}{\xi}\Big)C^2\end{aligned}$$

실험 $\mathcal{E} = \{P_\theta : \theta \in \Theta\}$가 $\mathcal{E}_j = \{P_{\theta,j} : \theta \in \Theta\}$ 의 direct product 일 때를 증명하려 한다. 여기서는 $\mathcal{E}$가 동일하지 않아도 된다.

먼저 Θ의 거리화를 위해 다음을 정의하자.

$$H^2(s,t) = \sum_j h^2(p_{s,j}, p_{t,j}) = \frac{1}{2}\sum_j \int (\sqrt{dP_{s,j}} - \sqrt{dP_{t,j}})^2.$$

(s,t) 쌍을 생각하자. $\xi \in (0, \frac{1}{3}]$이 되도록 고르면 중심이 $p_{s,j}$ 인 헬링거 공을 다음과 같이 정의할 수 있다.

$$A_j = \{p : h^2(p, p_{s,j}) < 1 - \cos \xi w_j\} \quad \text{where } w_j \in (0, \frac{\pi}{2}).$$

B_j 도 중심이 $p_{t,j}$가 되도록 정의할 수 있다.

레마1과 정리1 에 의해 다음을 만족하는 함수 ϕ_j가 존재한다.

$$\int \phi_j dp \le \cos(1-2\xi)w_j \quad \text{for} \quad \forall p \in A_j$$
$$\int \phi_j^{-1} dq \le \cos(1-2\xi)w_j \quad \text{for} \quad \forall q \in B_j$$

여기서 검정 함수 ψ는 다음과 같이 놓을 수 있다.

$$\psi = \begin{cases} 0, & \prod_j \phi_j < 1 \\ 1, & \prod_j \phi_j > 1 \\ \text{arbirary on } [0,1], & \prod_j \phi_j = 1 \end{cases}$$

레마 3 (레마2?,Prop3?). *Let $\mathcal{E} = \{P_\theta : \theta \in \Theta\}$ be a direct product of experiments $\mathcal{E}_j = \{P_{\theta,j} : \theta \in \Theta\}$, $j \in J$. Take a number $\xi \in (0, \frac{1}{3}]$ and let ψ be the test described before the statement of this lemma. Then for any number q such that $0 < q < 1 - 4\xi$, there is a number $\gamma > 0$ such that $\int \psi dP_\theta \leq \exp\big(-qH^2(s,t)\big)$ whenever $H(\theta, s) \leq \gamma H(s,t)$ and $\int (1-\psi) dP_\theta \leq \exp\big(-qH^2(s,t)\big)$ whenever $H(\theta,t) \leq \gamma H(s,t)$.*

In addition, there are choices of ξ such that $H(\theta', s) \leq \frac{1}{18}H(s,t)$ and $H(\theta'', t) \leq \frac{1}{18}H(s,t)$ imply, respectively,

$$\int \psi dP_{\theta'} \leq \exp\Big(-\frac{1}{2}H^2(s,t)\Big),$$

$$\&$$

$$\int (1-\psi) dP_{\theta''} \leq \exp\Big(-\frac{1}{2}H^2(s,t)\Big).$$

Proof. *Let s and t be two elements of Θ. 레마2의 (P,Q) 역할을 $(p_{s,j}, p_{t,j})$가 할 예정이다.*

$$\begin{aligned} h_j^2 &= h^2(p_{s,j}, p_{t,j}) = 2\sin^2 w_j/2 \\ r_j &= \sqrt{2}\sin \xi w_j/2 \quad \textit{for some} \xi \in (0, \frac{1}{3}]. \end{aligned}$$

이 때, 만약 $c_j = h(p_{\theta,j}, p_{s,j})$라고 하면, 레마2 에 의해 다음이 성립한다. ($H(\theta,s) \leq \gamma H(s,t)$ 를 가정하자.)

$$\begin{aligned} \int \prod_j \phi_j dP_\theta &\leq \prod_j \Big[1 - (1-2\xi)^2 h_j^2 + 4(\frac{1-2\xi}{\xi})c_j^2\Big] \\ &\leq \exp\Big(-(1-2\xi)^2 \sum h_j^2 + 4(\frac{1-2\xi}{\xi})\sum c_j^2\Big) \\ &\leq \exp\Big(-(1-2\xi)^2 \sum h_j^2 + 4(\frac{1-2\xi}{\xi})\gamma^2 \sum h_j^2\Big) \\ &= \exp\Big(-\big((1-2\xi)^2 - 4(\frac{1-2\xi}{\xi})\gamma^2\big)H^2(s,t)\Big) \\ &\leq \exp\Big(-\frac{1}{2}H^2(s,t)\Big) \quad (\textit{where}\, \gamma = \frac{1}{18}) \end{aligned}$$

이를 정리하면 다음과 같은 결론을 도출할 수 있다.

$$\int \psi dP_\theta \leq \int \prod_j \phi_j dP_\theta \leq \exp\Big(-\frac{1}{2}H^2(s,t)\Big) \quad \textit{where} \quad H(\theta,s) \leq \frac{1}{18}H(s,t).$$

비슷한 방법으로 다음의 결과도 얻을 수 있다.

$$\int (1-\psi) dP_\theta \le \int \prod_j \phi_j^{-1} dP_\theta \le \exp\Big(-\frac{1}{2}H^2(s,t)\Big) \quad where \quad H(\theta,t) \le \frac{1}{18}H(s,t).$$

2 레마3

레마 4 (Lemma 3 **?**)**.** *Assume that the following model,*

$$y_i = f_0(X_i) + \xi_i, \quad \xi_i \overset{iid}{\sim} N(0,\sigma^2) \quad i = 1, \cdots, n.$$

Suppose that $\mathcal{F}$ is uniformly bounded. Let

$$A_{\epsilon,M} := \{f \in \mathcal{F} : ||f - f_0||_n < M\epsilon\}.$$

If there exist $C > 2/\sigma^2$ and $\mathcal{F}_n \subset \mathcal{F}$ such that

$$\sup_{\epsilon > \epsilon_n} \log N(\epsilon/36, A_{\epsilon,1} \cap \mathcal{F}_n, ||\cdot||_n) \le n\epsilon_n^2, \tag{5}$$

$$\Pi(A_{\epsilon_n,1}) \ge e^{-Cn\epsilon_n^2}, \tag{6}$$

$$\Pi(\mathcal{F} - \mathcal{F}_n) = o(e^{-(C\sigma^2+2)n\epsilon_n^2}) \tag{7}$$

Then for any $\epsilon_n \to 0$, $n\epsilon_n^2 \to \infty$,

$$\Pi(A^c_{\epsilon_n,1M_n}|\mathbb{D}) \to 0$$

in $P_{f_0}^{(n)}$-probability as $n \to \infty$ for any $M_n \to \infty$.

Proof. $\bar{d}_n(f_0, f) = ||f_0 - f||_n$ 이라고 하자. **?** 의 214쪽에서도 언급되었듯이 헬링거 거리 대신에 $\bar{d}_n$를 사용하여도 된다.

$$K(P_{f_0,i}, P_{f,i}) = \frac{1}{2}V_{2,0}(P_{\theta_0,i}, P_{\theta,i}) = \frac{1}{2\sigma^2}|f_0(x_i) - f(x_i)|^2$$

이 된다. 따라서 $\bar{B}_n(f_0,\epsilon) = \{f \in \mathcal{F} : \bar{d}_n^2(f_0, f) \le \sigma^2\epsilon^2\} = A_{\sigma\epsilon_n,1}$ 이 된다. 레마 2를 이용하기 위해서는 충분히 큰 j에 대해서 다음 두 가지만 보이면 충분하다.

$$\frac{\Pi(\mathcal{F} - \mathcal{F}_n)}{A_{\sigma\epsilon_n,1}} = o(e^{-2n\epsilon_n^2}), \tag{8}$$

$$\frac{\Pi(A_{2j\epsilon_n,1} - A_{j\epsilon_n,1})}{A_{\sigma\epsilon_n,1}} \le e^{n\epsilon_n^2 j^2/4}. \tag{9}$$

식 (8) 은 다음과 같이 보일 수 있다.

$$\begin{aligned}\frac{\Pi(\mathcal{F} - \mathcal{F}_n)}{A_{\sigma\epsilon_n,1}} &\le e^{Cn\sigma^2\epsilon_n^2}\Pi(\mathcal{F} - \mathcal{F}_n) \quad (by\ (6)) \\ &= o(e^{-2n\epsilon_n^2}) \quad (by\ (7))\end{aligned}$$

식 (9) 은 다음과 같이 보일 수 있다.

$$\frac{\Pi(A_{2j\epsilon_n,1} - A_{j\epsilon_n,1})}{A_{\sigma\epsilon_n,1}} \leq e^{Cn\sigma^2\epsilon_n^2}\Pi(A_{2j\epsilon_n,1} - A_{j\epsilon_n,1}) \quad (by\ (6))$$
$$\leq e^{n\epsilon_n^2 j^2/4}$$

3 정리1

정리 2 (Theorem 1 **?**)**.** *Assume that the following model,*

$$y_i = f_0(X_i) + \xi_i, \quad \xi_i \overset{iid}{\sim} N(0, \sigma^2) \quad i = 1, \cdots, n,$$

and further assume the following prior distribution,

$$\pi(\theta_j|\gamma_j, L, W, S, B) = \gamma_j \tilde{\pi}(\theta_j|L, W, S, B) + (1 - \gamma_j)\delta_0(\theta_j),$$
$$\pi(\gamma|L, W, S, B) = 1/\begin{pmatrix} T \\ S \end{pmatrix},$$
$$\pi(L = L_n) = \pi(W = W_n) = \pi(S = S_n) = \pi(B = B_n) = 1.$$

where $\tilde{\pi}(\theta_j|L, W, S, B) = U(\theta_j; [-B, B])$ *and* $T = |\Theta(L, W, S, B)|$*. Suppose that* $0 < F < \infty$*,* $0 < p, q \leq \infty$ *and* $d(1/p - 1/2)_+ < s$*. Then the posterior distribution concentrates at the true function with a rate* $\epsilon_n = n^{-s/(2s+d)}(\log n)^{3/2}$*. That is,*

$$\Pi(f_\theta \in \Phi \cap UB(F) : ||f_\theta - f_0||_n > M_n\epsilon_n|\mathbb{D}_n) \to 0$$

in $P_{f_0}^{(n)}$*-probability as* $n \to \infty$ *for any* $M_n \to \infty$*.*

Proof. $\mathcal{F} = \Phi \cap UB(F)$ 라고 하자. 레마 4 에 의해서 다음을 만족하는 $C > 2/\sigma^2$ 와 $\mathcal{F}_n \subset \mathcal{F}$ 가 존재함을 보이면 충분하다.

$$\sup_{\epsilon > \epsilon_n} \log N(\epsilon/36, A_{\epsilon,1} \cap \mathcal{F}_n, ||\cdot||_n) \leq n\epsilon_n^2, \tag{10}$$
$$-\log \Pi(A_{\epsilon_n,1}) \leq -Cn\epsilon_n^2, \tag{11}$$
$$\Pi(\mathcal{F} - \mathcal{F}_n) = o(e^{-(C\sigma^2+2)n\epsilon_n^2}) \tag{12}$$

- $\mathcal{F}_n = \Phi(L_n, W_n, S_n, B_n) \cap UB(F)$ 라고 하면 사전분포에 의해서 식 (12)은 자명함을 알 수 있다.
- 다음으로는 식 (10)이 성립함을 보이려고 한다.

$$\{f \in \mathcal{F}_n : ||f||_{L^\infty} \leq \epsilon\} \subset \{f \in \mathcal{F}_n : ||f||_n \leq \epsilon\} \tag{13}$$

라는 사실과 레마4로부터 다음을 알 수 있다.

$$\sup_{\epsilon>\epsilon_n} \log N(\epsilon/36, A_{\epsilon,1}\cap\mathcal{F}_n, ||\cdot||_n) \tag{14}$$

$$\leq \sup_{\epsilon>\epsilon_n} \log N(\epsilon/36, A_{\epsilon,1}\cap\mathcal{F}_n, ||\cdot||_{L^\infty}) \quad (by\ (13)) \tag{15}$$

$$\leq \sup_{\epsilon>\epsilon_n} \log N(\epsilon/36, \mathcal{F}_n, ||\cdot||_{L^\infty}) \tag{16}$$

$$\leq \log N(\epsilon_n/36, \Phi(L_n, W_n, S_n, B_n), ||\cdot||_{L^\infty}) \tag{17}$$

$$\leq (S_n+1)\Big[\log L_n + L_n \log\big((B_n\vee 1)(W_n+1)^2\big) - \log\frac{\epsilon_n}{72}\Big] \quad (by\ \text{레마4}) \tag{18}$$

$$\lesssim N_n(\log n)^3 \tag{19}$$

$$= n\epsilon_n^2 \tag{20}$$

부등식 (19)은 다음과 같은 사실들에 의해서 성립한다.

$$N_n = \lceil nd/(2s+d)\rceil,\ W_0 = 6dm(m+2)+2d$$

$$L_n = L(N_n),\ W_n = N_n W_0$$

$$S_n = (L_n-1)W_0^2 N_n + N_n,\ B_n = O(N_n^{\Xi}),$$

$$\text{where}\quad c_{(d,m)} = (1+2de(2e)^m/\sqrt{m})^{-1},\ L(N_n) = 3+\lceil\log_2(\frac{3^{d\vee m}}{\tau_n c_{(d,m)}})+5\rceil\lceil\log_2(d\vee m)\rceil$$

$$\tau_n = N_n^{-s/d-(\nu^{-1}+d^{-1})(d/p-s)_+}(\log N_n)^{-1},\ \Xi = (\nu^{-1}+d^{-1})(1\vee(d/p-s)_+) \tag{21}$$

- 다음으로는 식 (11)이 성립함을 보이려고 한다. 레마 6에 의해서 다음이 성립한다.

$$\left|\left|\hat{f}_n - f_0\right|\right|_{L^2} \leq C_1||f_0||_{B^s_{p,q}([0,1]^d)} N_n^{-s/d} \leq \epsilon_n/4. \tag{22}$$

강대수의 법칙과 $p_X(x)\leq R\leq 2$의 조건으로 인해서 다음이 성립한다.

$$\left|\left|\hat{f}_n - f_0\right|\right|_n^2 \leq 2\left|\left|\hat{f}_n - f_0\right|\right|_{L^2(P_X)}^2 \leq 4\left|\left|\hat{f}_n - f_0\right|\right|_{L^2}^2. \tag{23}$$

이제 $\hat{\gamma}$와 $\hat{\theta}_{\hat{\gamma}}$ 를 0이 아닌 인덱스와 해당 인덱스에서의 값으로 정의하자. $\Theta(\hat{\gamma}, L_n, W_n, S_n, B_n)$ 은 $\Theta(L_n, W_n, S_n, B_n)$ 에서 $\hat{\gamma}$ 부분과 동일한 0이 아닌 인덱스를 가지는 부분집합을 의미한다. 또한 $\mathcal{F}_n(\hat{\gamma}) = \Phi(L_n, W_n, S_n, B_n)\cap UB(F)$ 로 정의하면 다음을 보일 수 있다.

$$\begin{aligned}
\Pi(A_{\epsilon_n,1}) &= \Pi(f\in\mathcal{F}_n : ||f-f_0||_n \leq \epsilon_n)\\
&\geq \Pi(f\in\mathcal{F}_n : ||f-f_0||_{L^2} \leq \epsilon_n/2) \quad (by\ (23))\\
&\geq \Pi(f\in\mathcal{F}_n : \left|\left|f-\hat{f}_n\right|\right|_{L^2} \leq \epsilon_n/4) \quad (by\ (22))\\
&\geq \Pi(f\in\mathcal{F}_n : \left|\left|f-\hat{f}_n\right|\right|_{L^\infty} \leq \epsilon_n/4) \quad (by\ (13))\\
&\geq \Pi(f\in\mathcal{F}_n(\hat{\gamma}) : \left|\left|f-\hat{f}_n\right|\right|_{L^\infty} \leq \epsilon_n/4)
\end{aligned}$$

레마4 의 증명과정에서 다음과 같은 사실을 알 수 있다.

$$||\theta - \theta^*||_\infty < \epsilon \Longrightarrow |f_\theta(x) - f_{\theta^*}(x)| \leq \epsilon L(B \vee 1)^{L-1}(W+1)^L \tag{24}$$

$$\binom{T}{S} \leq (W+1)^{Ls} \tag{25}$$

이로부터 다음과 같은 식을 얻을 수 있다.

$$\begin{aligned}
&\Pi(f \in \mathcal{F}_n(\hat{\gamma}) : \left|\left|f - \hat{f}_n\right|\right|_{L^\infty} \leq \epsilon_n/4) \\
&\geq \Pi(\theta \in \mathbb{R}^{T_n} : \theta_{\hat{\gamma}^c} = 0,\ ||\theta||_\infty \leq B_n,\ ||\theta - \theta^*||_\infty \leq \frac{\epsilon_n}{4L_n(B_n \vee 1)^{L_n-1}(W_n+1)^{L_n}}) \quad (by\ (24)) \\
&\geq \left(\frac{\epsilon_n}{4B_nL_n(B_n \vee 1)^{L_n-1}(W_n+1)^{L_n}}\right)^{S_n} \binom{T}{S}^{-1} \qquad \text{(by prior setting)} \\
&\geq \left(\frac{\epsilon_n}{4B_nL_n(B_n \vee 1)^{L_n-1}(W_n+1)^{2L_n}}\right)^{S_n} \qquad (by\ (25)) \\
&= \exp\Big(-S_n \log\Big(\frac{4B_nL_n(B_n \vee 1)^{L_n-1}(W_n+1)^{2L_n}}{\epsilon_n}\Big)\Big)
\end{aligned}$$

따라서 이를 정리하면 다음과 같이 된다.

$$\begin{aligned}
-\log \Pi(A_{\epsilon_n,1}) &\leq S_n \log\Big(\frac{4B_nL_n(B_n \vee 1)^{L_n-1}(W_n+1)^{2L_n}}{\epsilon_n}\Big) \\
&\leq S_n[L_n \log\big((W_n+1)^2(B_n \vee 1)\big) + \log 4L_n - \log \epsilon_n] \\
&\lesssim N_n(\log n)^3 \quad (by\ (21) \\
&= n\epsilon_n^2
\end{aligned}$$