نظریه اطلاعات، آمار و یادگیری

دکتر یاسایی

دانشگاه صنعتی شریف

مهندسی برق

برنا خدابنده ۴۰۰۱۰۹۸۹۸

تمرین شماره Final - Q1

تاریخ: 1403/03/05

Problem 1) Variational form χ^2: $\sup_{h:\mathcal{X}\to\mathbb{R}} \left\{ \mathbb{E}_P[h(x)] - \mathbb{E}_Q\left[\frac{h^2(x)}{4}\right] - 1 \right\} = \chi^2(P\|Q)$

Since: $D_f(P\|Q) = \sup_h \left\{ \mathbb{E}_P[h(x)] - \mathbb{E}_Q[f^*(h(x))] \right\}$, χ^2: $f(x) = x^2 - 1 \rightsquigarrow f^*(y) = \frac{1}{4}y^2 + 1$

let: $h(x) = f(x) + c$, $c \in \mathbb{R}$, $f: \mathcal{X} \to \mathbb{R} \rightsquigarrow \sup_{h:\mathcal{X}\to\mathbb{R}} \{\dots\} = \sup_{f:\mathcal{X}\to\mathbb{R}} \sup_{c\in\mathbb{R}} \{\dots\}$

$$\chi^2(P\|Q) = \sup_{h:\mathcal{X}\to\mathbb{R}} \{\dots\} = \sup_{f:\mathcal{X}\to\mathbb{R}} \sup_{c\in\mathbb{R}} \left\{ \mathbb{E}_P[f(x)+c] - \mathbb{E}_Q\left[\frac{1}{4}(f(x)+c)^2\right] - 1 \right\}$$

$$= \sup_{f:\mathcal{X}\to\mathbb{R}} \sup_{c\in\mathbb{R}} \Big\{ \underbrace{\mathbb{E}_P[f(x)] + c - \frac{1}{4}\mathbb{E}_Q[f^2(x)] - \frac{1}{2}\mathbb{E}_Q[f(x)]c - \frac{1}{4}c^2 - 1}_{\mathcal{L}(x;c,f)} \Big\}$$

@ $c_0 = \arg\max \mathcal{L}(x;c,f)$: $\frac{\partial \mathcal{L}(x;c,f)}{\partial c} = 0 \rightsquigarrow 1 - \frac{1}{2}\mathbb{E}_Q[f(x)] - \frac{1}{2}c_0 = 0$

$\Rightarrow c_0 = 2 - \mathbb{E}_Q[f(x)] \rightsquigarrow \chi^2 = \sup_{f:\mathcal{X}\to\mathbb{R}} \mathcal{L}(x;c_0,f)$

$$\chi^2(P\|Q) = \sup_{f:\mathcal{X}\to\mathbb{R}} \Big\{ \mathbb{E}_P[f(x)] + 2 - \mathbb{E}_Q[f(x)] - \frac{1}{4}\mathbb{E}_Q[f^2(x)] - 1 - \mathbb{E}_Q[f(x)] + \frac{1}{2}\mathbb{E}_Q[f(x)]^2 \underbrace{- \frac{1}{4}(2 - \mathbb{E}_Q[f(x)])^2}_{-1 + \mathbb{E}_Q[f(x)] - \frac{1}{4}\mathbb{E}_Q[f(x)]^2} \Big\} = \sup_{f:\mathcal{X}\to\mathbb{R}} \left\{ \mathbb{E}_P[f(x)] - \mathbb{E}_Q[f(x)] + \frac{1}{4}\left(\mathbb{E}_Q[f(x)]^2 - \mathbb{E}_Q[f^2(x)]\right) \right\}$$

$$\Rightarrow \boxed{\chi^2(P\|Q) = \sup_{f:\mathcal{X}\to\mathbb{R}} \left\{ \mathbb{E}_P[f(x)] - \mathbb{E}_Q[f(x)] - \frac{1}{4}\mathrm{Var}_Q[f(x)] \right\}}$$

$$\chi^2(P_{XY}\|Q_XQ_Y) = \sup_{f:\mathcal{X}\times\mathcal{Y}\to\mathbb{R}} \left\{ \mathbb{E}_P[f(x,y)] - \mathbb{E}_Q[f(x,y)] - \frac{1}{4}\mathrm{Var}_Q[f(x,y)] \right\}$$

$\to Q_{XY} = Q_XQ_Y$
X, Y are independent

sum of variances

$$\geq \sup_{\substack{f(x,y) = h(x) + g(y) \\ h:\mathcal{X}\to\mathbb{R} \\ g:\mathcal{Y}\to\mathbb{R}}} \{\dots\} = \sup_{g:\mathcal{Y}\to\mathbb{R}} \sup_{h:\mathcal{X}\to\mathbb{R}} \Big\{ \mathbb{E}_{P_{XY}}[h(x)+g(y)] - \mathbb{E}_{Q_XQ_Y}[h(x)+g(y)] - \frac{1}{4}\underbrace{\mathrm{Var}_{Q_XQ_Y}[h(x)+g(y)]}_{\text{by independance}} \Big\}$$

$$= \sup_{g:\mathcal{Y}\to\mathbb{R}} \sup_{h:\mathcal{X}\to\mathbb{R}} \Big\{ \mathbb{E}_{\cancel{P_{XY}}\,P_X}[h(x)] + \mathbb{E}_{\cancel{P_{XY}}\,P_Y}[g(y)] - \mathbb{E}_{\cancel{Q_XQ_Y}\,Q_X}[h(x)] - \mathbb{E}_{\cancel{Q_XQ_Y}\,Q_Y}[g(y)] - \frac{1}{4}\left(\mathrm{Var}_{Q_X}(h(x)) + \mathrm{Var}_{Q_Y}(g(y))\right) \Big\}$$

$$= \sup_{g:\mathcal{Y}\to\mathbb{R}} \sup_{h:\mathcal{X}\to\mathbb{R}} \Big\{ \overbrace{\left(\mathbb{E}_{P_X}[h(x)] - \mathbb{E}_{Q_X}[h(x)] - \frac{1}{4}\mathrm{Var}_{Q_X}[h(x)]\right)}^{\chi^2(P_X\|Q_X)} + \overbrace{\left(\mathbb{E}_{P_Y}[g(y)] - \mathbb{E}_{Q_Y}[g(y)] - \frac{1}{4}\mathrm{Var}_{Q_Y}[g(y)]\right)}^{\chi^2(P_Y\|Q_Y)} \Big\}$$

$$\Longrightarrow \boxed{\chi^2(P_{XY}\|Q_XQ_Y) \geq \chi^2(P_X\|Q_X) + \chi^2(P_Y\|Q_Y)}$$