Секция «Математика и механика»

О приращении f-дивергенции

Клиндухов Дмитрий Владимирович
Студент
Московский государственный университет имени М.В. Ломоносова, Механико-математический факультет, Москва, Россия
E-mail: dmi31992@gmail.com

В докладе рассматривается следующая задача.
Пусть даны два бинарных эксперимента $\mathbb{E}_0 = \left(\Omega_0, \mathcal{F}_0, \left(P_0, \tilde{P}_0\right)\right)$ и $\mathbb{E}_1 = \left(\Omega_1, \mathcal{F}_1, \left(P_1, \tilde{P}_1\right)\right)$, причем известно, что $\mathbb{E}_1$ более информативен, чем $\mathbb{E}_0$. Кроме того, предполагается, что $\|P_0 - \tilde{P}_0\| = 2v_0$ и $\|P_1 - \tilde{P}_1\| = 2v_1$, где v_0 и v_1 фиксированы. Требуется найти минимум $\left|J_f\left(P_1, \tilde{P}_1\right) - J_f\left(P_0, \tilde{P}_0\right)\right|$ по всем таким экспериментам.
Эквивалентная постановка задачи получается, если предположить, что $\Omega_0 = \Omega_1$, $\mathcal{F}_0 \subseteq \mathcal{F}_1$, $P_0 = P_1|_{\mathcal{F}_0}$, $\tilde{P}_0 = \tilde{P}_1|_{\mathcal{F}_0}$.
Для точной формулировки результата введем необходимые определения. Эксперимент $\mathbb{E}_1 = \left(\Omega_1, \mathcal{F}_1, \left(P_1, \tilde{P}_1\right)\right)$ более информативен, чем эксперимент $\mathbb{E}_0 = \left(\Omega_0, \mathcal{F}_0, \left(P_0, \tilde{P}_0\right)\right)$, если для любой измеримой $\varphi_0 : (\Omega_0; \mathcal{F}_0) \to [0; 1]$ найдется такая измеримая функция $\varphi_1 : (\Omega_1; \mathcal{F}_1) \to [0; 1]$, что $\int \varphi_1 dP_1 \leq \int \varphi_0 dP_0$ и $\int (1 - \varphi_1)\, dP_1 \leq \int (1 - \varphi_0)\, dP_0$, т.е. для любого теста в эксперименте $\mathbb{E}_1$ найдется тест в эксперименте $\mathbb{E}_0$, у которого ошибки первого и второго рода не больше, чем соответствующие ошибки у первого теста.
Понятие f-дивергенции вероятностных мер было введено И.Чисаром в 1963 году и включает в себя многие известные расстояния. Пусть $f : (0; +\infty) \to \mathbb{R}$ - выпуклая функция. Положим $f(0) = \lim\limits_{u \to 0+} f(u)$ и

$$F(u, v) = \begin{cases} 0 & \text{если } u = 0 \text{ и } v = 0, \\ v f(u/v) & \text{если } u \geq 0 \text{ и } v > 0, \\ u \lim\limits_{u \to \infty} \frac{f(u)}{u} & \text{если } u > 0 \text{ и } v = 0, \end{cases}$$

f-дивергенцией вероятностных мер P и Q, заданных на одном вероятностном пространстве, называется величина $J_f := \int F\left(\frac{dP}{d\rho}, \frac{dQ}{d\rho}\right) d\rho$, где ρ - произвольная сигма-аддитивная мера на $(\Omega, \mathcal{F})$, такая что $P \ll \rho$ и $Q \ll \rho$.
Известно [1], что $J_f(P, Q) = J_f(Q, P)$ для любых P и Q тогда и только тогда, когда $f(u) = u f\left(\frac{1}{u}\right)$ для любого $u > 0$. Частными случаями f-дивергенции является расстояние по вариации $\|P_0 - \tilde{P}_0\| := 2 \sup\limits_{A \in \mathcal{F}} |P(A) - Q(A)| = J_f(P, Q)$, где функция $f(u) = |u - 1|$, квадрат расстояния Хеллингера $\rho^2(P, Q) = J_f(P, Q)$, где $f(u) = \frac{1}{2}\left(\sqrt{u} - 1\right)^2$, и другие.
Эксперимент $E_1 = \left(\Omega_1, \mathcal{F}_1, \left(P_1, \tilde{P}_1\right)\right)$ более информативен, чем эксперимент $E_0 = \left(\Omega_0, \mathcal{F}_0, \left(P_0, \tilde{P}_0\right)\right)$, тогда и только тогда, когда $J_f(P_1, \tilde{P}_1) \geq J_f(P_0, \tilde{P}_0)$ для любой выпуклой f.
Теорема. Пусть $\mathbb{E}_0$ и $\mathbb{E}_1$ - два бинарных эксперимента, $\mathbb{E}_1$ более информативен, чем $\mathbb{E}_0$, $\|P_0 - \tilde{P}_0\| = 2v_0$, $\|P_1 - \tilde{P}_1\| = 2v_1$. Предположим также, что $f : (0; +\infty) \to \mathbb{R}$ -

выпуклая функция и $f(u) = uf\left(\frac{1}{u}\right)$ для любого $u > 0$. Тогда, в предположении, что $J_f\left(P_0, \tilde{P}_0\right) < \infty$,

$$\left|J_f\left(P_0, \tilde{P}_0\right) - J_f\left(P_1, \tilde{P}_1\right)\right| \geq (1 - v_1)\left(f\left(\frac{1 + v_1 - 2v_0}{1 - v_1}\right) - f(1)\right)$$

причем для любых $0 \leq v_0 \leq v_1 \leq 1$ можно построить эксперименты с заданными условиями, на которых достигается равенство.
Частными случаями теоремы являются хорошо известные результаты (см. Liese, Vajda [1, глава 8]) о минимуме и максимуме f-дивергенции при заданном расстоянии по вариации. Минимальное значение отвечает случаю, когда $v_0 = 0$, т.е. $P_0 = \tilde{P}_0$. Максимальное - случаю, когда $v_1 = 1$, т.е. P_1 и $\tilde{P}_1$ сингулярны.

Литература

1. Liese F., Vajda I. Convex statistical distances. Leipzig, 1987.

Слова благодарности

Автор благодарен своему научному руководителю А.А. Гущину за постановку задачи и внимание к работе.