О некоторых свойствах процедур рандомизированного машинного обучения при наличии зашумленных данных*

Ю. С. Попков

Федеральный исследовательский центр «Информатика и управление» РАН, Москва, Россия

Аннотация. Рассматриваются различные модели измерительных шумов в процедурах рандомизированного энтропийного оценивания функций плотности распределения вероятностей: аддитивные, мультипликативные, измерительные шумы на входе и выходе модели объекта. Исследуются свойства энтропийно-оптимальных ПРВ, и показано, что соответствующие им измерительные шумы являются гетероскедастическими.

Ключевые слова: энтропийное оценивание, функции плотности, множители Лагранжа, гетероскедастический шум, модели дисперсии.

DOI 10.14357/20718632230209

Введение

Процедуры рандомизированного машинного обучения являются эффективным средством решения задач классификации, прогнозирования и кластеризации в условиях неопределенности, прежде всего, данных [1]. Особенность этих процедур в том, что решение указанных задач осуществляется в терминах двух типов ансамблей случайных ответов: энтропийно оптимальных параметризованных моделей и измерительных шумов с максимальной энтропийной неопределенностью [2-5].

В [6] показано, что оценки максимальной энтропии функции плотности распределения вероятностей параметров моделей являются асимптотически эффективными, т.е. стремятся к стационарным формам при расширении объема наблюдений.

Настоящая работа посвящена исследованию свойств энтропийно-оптимальных измерительных шумов при аддитивных и мультипликативных моделях наблюдений. Оптимальные функции плотности распределения вероятностей измерительных шумов соответствуют максимальной неопределенности, т.е. в каком-то (энтропийном) смысле являются наихудшими. Более глубокое их исследование показало, что энтропийно-оптимальные измерительные шумы имеют гетероскедастическую природу.

1. Наблюдения с аддитивным измерительным шумом

Рассмотрим случай обучения нелинейной статической модели со случайными, интервальными параметрами:

$$y[n] = \varphi(f[n], a), \quad a \in \mathcal{A} = [a^-, a^+], \tag{1}$$

* Работа поддержана Министерством науки и высшего образования РФ, проект № 075-15-2020-799.

характеризуемыми функцией ПРВ $P(a)$. Наблюдаемый выход модели

$$v[n] = y[n] + \xi[n], \quad \xi[n] \in \Xi_n = [\xi_n^-, \xi_n^+]. \tag{2}$$

Для обучения используются данные

$$y^r[1], \dots, y^r[N], \tag{3}$$

на интервале $\mathcal{N} = [1, N]$.

Измерительный шум предполагается независимым с функцией ПРВ

$$Q(\xi[1], \dots, \xi[N]) = \prod_{n=1}^{N} Q_n(\xi[n]) \tag{4}$$

Функции ПРВ параметров и шумов определяются решением следующей задачи:

$$H(P(a), Q(\xi[1], \dots, \xi[N])) = -\int_{\mathcal{A}} P(a) \ln P(a)\, da - \\ -\sum_{n=1}^{N} \int_{\Xi_n} Q_n(\xi[n]) \ln Q_n(\xi[n])\, d\xi[n] \Rightarrow \max, \tag{5}$$

$$\int_{\mathcal{A}} P(a)\, da = 1, \quad \int_{\Xi_n} Q_n(\xi[n])\, d\xi[n] = 1, \; n = \overline{1, N},$$

$$\int_{\mathcal{A}} P(a)\, \varphi(f[n], a) da + \int_{\Xi_n} Q_n(\xi[n])\, \xi[n]\, d\xi[n] = y^r[n], \; n = \overline{1, N}. \tag{6}$$

Теорема 1. *Энтропийно оптимальные измерительные шумы в задаче имеют переменную дисперсию, т.е. являются гетероскедастичными.*

Доказательство. Согласно [1]

$$Q_n^*(\xi[n]) = \frac{\exp(-\lambda_n\, \xi[n])}{\mathcal{Q}_n(\lambda_n)}, \quad \mathcal{Q}_n(\lambda_n) = \int_{\Xi_n} \exp(-\lambda_n\, \xi[n]), \; n = \overline{1, N}, \tag{7}$$

где λ_n - множители Лагранжа, определяемые уравнениями эмпирических балансов:

$$\mathcal{P}^{-1}(\boldsymbol{\lambda}) \int_{\mathcal{A}} \exp\left(-\sum_{i=1}^{N} \lambda_i\, \varphi(f[i], a)\right) \varphi(f[n], a)\, da + \\ +\mathcal{Q}_n^{-1}(\lambda_n) \int_{\Xi_n} \exp(-\lambda_n \xi[n])\, \xi[n]\, d\xi[n] = y^r[n], \; n = \overline{1, N}. \tag{8}$$

По определению, среднее значение интервального шум:а

$$\bar{\xi}[n] = \int_{\Xi_n} \xi[n]\, \frac{\exp(-\lambda_n\, \xi[n])}{\mathcal{Q}_n(\lambda_n)}\, d\xi[n] = \bar{\xi}[n](\lambda_n), \tag{9}$$

и его дисперсия:

$$D_{\xi_{[n]}} = \int_{\Xi_n} \frac{\exp(-\lambda_n\, \xi[n])}{\mathcal{Q}_n(\lambda_n)} \left(\xi[n] - \bar{\xi}[n](\lambda_n)\right)^2 d\xi[n] = D_{\xi[n]}(\lambda_n). \tag{10}$$

Из этих равенств следует утверждение теоремы 1.

2. Наблюдения с мультипликативным измерительным шумом

Рассмотрим модель и мультипликативные наблюдения

$$y[n] = \varphi(f[n], a), \quad v[n] = \varphi(f[n], a)\, \xi[n], \; n = \overline{1, N}. \tag{11}$$

Параметры a и шум $\boldsymbol{\xi} = \{\xi[1], \dots, \xi[N]\}$ случайные, независимые, интервальные (1, 2).

Введем совместную функцию ПРВ параметров и шумов $F(a, \boldsymbol{\xi})$, определенную на множестве

$$\mathcal{F} = \mathcal{A} \otimes \left(\bigotimes_{n=1}^{N} \Xi_n\right). \tag{12}$$

Для оценивания функции ПРВ $F(a, \boldsymbol{\xi})$ воспользуемся алгоритмом РМО:

$$\mathcal{H}[F(a, \boldsymbol{\xi})] = -\int_{\mathcal{F}}) F(a, \boldsymbol{\xi}) \ln F(a, \boldsymbol{\xi}) da\, d\boldsymbol{\xi} \Rightarrow \max, \tag{13}$$

$$\int_{\mathcal{F}} F(a, \boldsymbol{\xi})\, da\, d\boldsymbol{\xi} = 1, \tag{14}$$

$$\int_{\mathcal{F}} F(a, \boldsymbol{\xi})\, \varphi(f[n], a)\, \xi[n]\, da\, d\boldsymbol{\xi} = y^r[n], \quad n = \overline{1, N}. \tag{15}$$

Теорема 2. *Энтропийно оптимальные измерительные шумы в задаче (13-15) имеют переменную дисперсию, т.е. являются гетероскедастичными.*

Доказательство. Воспользуемся схемой доказательства теоремы 1 [1], адаптируя ее к терминам данной задаче.

Определим функционал Лагранжа

$$L[F(a,\boldsymbol{\xi}) \mid \mu, \boldsymbol{\lambda}] = \mathcal{H}[F(a,\boldsymbol{\xi})] + \mu\big(1 - \int_{\mathcal{F}} F(a,\boldsymbol{\xi}) da\, d\boldsymbol{\xi}\big) + \\ + \sum_{n=1}^{N} \lambda_n \big(y^r[n] - \int_{\mathcal{F}} F(a,\boldsymbol{\xi})\, \varphi(f[n],a)\, \xi[n]\, da\, d\xi[n]\big). \quad (16)$$

Условия оптимальности в терминах производных Гато определяют энтропийно-оптимальную функцию ПРВ

$$F^*(a,\boldsymbol{\xi} \mid \boldsymbol{\lambda}) = \prod_{n=1}^{N} \phi_n^*(a,\xi[n] \mid \lambda_n),$$
$$\phi_n^*(a,\xi[n] \mid \lambda_n) = \frac{\exp(-\lambda_n\, \varphi(f[n],a)\, \xi[n])}{\psi_n(\lambda_n)}, \quad (17)$$
$$\psi_n(\lambda_n) = \big(\int_{\mathcal{F}} \exp(-\sum_{n=1}^{N} \lambda_n\, \varphi(f[n],a)\, \xi[n])\, da\, d\boldsymbol{\xi}\big)^{\frac{1}{N}}.$$

Множители Лагранжа $\boldsymbol{\lambda}$ определяются балансовыми уравнениями (15). Из равенств (17) видно, что энтропийно-оптимальная функция ПРВ параметров и шумов существенно зависит от наблюдений $y^r[1], \dots, y^r[N]$ и от вида модели (1).

Среднее значение измерительного шума в момент времени n равно

$$\bar{\xi}_n = \int_{\mathcal{A}} da \int_{\Xi_n} \xi[n]\, \phi_n^*(a,\xi[n] \mid \lambda_n)\, d\xi[n], \quad (18)$$

и его дисперсия

$$D_{\xi[n]} = \int_{\mathcal{A}} da \int_{\Xi_n} (\xi[n] - \bar{\xi}_n)^2\, \phi_n^*(a,\xi[n] \mid \lambda_n)\, d\xi[n] = D_{\xi[n]}(\lambda_n). \quad (19)$$

Из этих равенств следует утверждение теоремы 2.

3. Модели зависимостей с измерительным шумами на входе и выходе

Рассмотрим модель зависимости следующего вида:

$$y[n] = \varphi(u[n]\ a), \quad u[n] = f[n] + \zeta[n], \quad v[n] = y[n] + \xi[n], \ n = \overline{1,N}. \quad (20)$$

Функция $\varphi(u[n],a)$ - непрерывная и представима полиномом степени R:

$$\varphi(u[n]\ a) = \sum_{k=1}^{R} a_k\, u^k[n] = \sum_{k=1}^{R} a_k \sum_{i=0}^{k} C_k^i\, f^{(k-i)}[n]\, \zeta^i[n]. \quad (21)$$

Здесь $\zeta[n]$ и $\xi[n]$ - измерительные шумы на входе и выходе соответственно. Они предполагаются случайными, независимыми и интервальными:

$$\zeta[n] \in \mathcal{Z}_n = [\zeta_n^-, \zeta_n^+], \quad \xi[n] \in \Xi_n = [\xi_n^-, \xi_n^+]. \quad (22)$$

Параметры модели (20) также случайные, независимые и интервальные:

$$a \in \mathcal{A} = [a^-, a^+]. \quad (23)$$

Вероятностные свойства случайных переменных будем характеризовать совместной функцией ПРВ $F(a,\boldsymbol{\zeta})$ и функциями ПРВ $Q_n(\xi[n]), n = \overline{1,N}$, где $\boldsymbol{\zeta} = \{\zeta[1], \dots, \zeta[N]\}$. Область определения функции $F(a,\boldsymbol{\zeta})$

$$\mathcal{F} = \mathcal{A} \otimes \Big(\bigotimes_{n=1}^{N} \mathcal{Z}_n\Big). \quad (24)$$

Для оценивания указанных ПРВ воспользуемся алгоритмом РМО [1]:

$$\mathcal{H}[F(a,\boldsymbol{\zeta}), Q(\boldsymbol{\xi})] = -\int_{\mathcal{F}} F(a,\boldsymbol{\zeta}) \ln F(a,\boldsymbol{\zeta})\, da\, d\boldsymbol{\zeta} - \\ - \sum_{n=1}^{N} \int_{\Xi_n} Q_n(\xi[n]) \ln Q_n(\xi[n])\, d\xi[n] \Rightarrow \max; \quad (25)$$

при ограничениях:

- нормировки

$$\int_{\mathcal{F}} F(a,\boldsymbol{\zeta})\, da\, d\boldsymbol{\zeta} = 1, \quad \int_{\Xi_n} d\xi[n] = 1, \; n = \overline{1,N}; \tag{26}$$

- эмпирических балансов

$$\int_{\mathcal{F}} F(a,\boldsymbol{\zeta}) \sum_{k=1}^{R} a_k \sum_{i=0}^{k} C_k^i f^{(k-i)}[n]\, \zeta^i[n]\, da\, d\boldsymbol{\zeta} + \\ + \sum_{n=1}^{N} \int_{\Xi_n} \xi[n]\, Q_n(\xi[n])\, d\xi[n] = y^r[n], \; n = \overline{1,N}. \tag{27}$$

Решение этой задачи имеет вид:

$$F^*(a,\boldsymbol{\zeta}) = \frac{\exp\left(-\sum_{n=1}^{N} \lambda_n \sum_{k=1}^{R} a_k \sum_{i=0}^{k} C_k^i f^{(k-i)}[n]\, \zeta^i[n]\right)}{\mathbb{F}(\boldsymbol{\lambda})}, \\ Q_n^*(\xi[n]) = \frac{\exp(-\lambda_n \xi[n])}{\mathbb{Q}_n(\lambda_n)}, \quad n = \overline{1,N}. \tag{28}$$

Здесь нормировочные функции имеют вид:

$$\mathbb{F}(\boldsymbol{\lambda}) = \int \mathcal{F} \exp\left(-\sum_{n=1}^{N} \lambda_n \sum_{k=1}^{R} a_k \sum_{i=0}^{k} C_k^i f^{(k-i)}[n]\, \zeta^i[n]\right) da\, d\boldsymbol{\zeta}, \\ \mathbb{Q}_n(\lambda_n) = \int_{\Xi_n} \exp(-\lambda_n \xi[n])\, d\xi[n], \quad n = \overline{1,N}. \tag{29}$$

Множители Лагранжа $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_N\}$ определяются решением балансовых уравнений (27).

Утверждения теорем 1 и 2 справедливы и в данном случае.

4. Восстановление закона изменения гетероскедастического шума

В предыдущем разделе было показано, что энтропийно-оптимальный измерительный интервальный шум имеет переменную в каждом измерительном акте дисперсию. Но ее величина ограничена размером интервала, в котором лежат значения измерительного шума.

Но бывают ситуации, когда изменение дисперсии гетероскедастического шума может оказаться произвольной. Для исследования характера этих изменений воспользуемся некоторыми видами математических моделей, в которых выделены случайная интервальная составляющая и составляющая, характеризующая изменение дисперсии, и они входят в модель мультипликативно [7].

Рассмотрим случай, довольно часто обсуждаемый в публикациях, посвященных гетероскедастичности [8, 9], когда предполагается, что дисперсия σ измерительных ошибок изменяется во времени. Обычно, рассматривается полиномиальная, положительная модель этой зависимости. В рамках концепции рандомизации будем рассматривать полиномиальную модель со случайными, интервальными параметрами:

$$\sigma(n,b) = \sum_{k=0}^{s} b_k n^k, \quad b_k \in \mathcal{B}_k = [0, b_k^+], \; \mathcal{B} = \bigotimes_{k=0}^{s} \mathcal{B}_k. \tag{30}$$

Вероятностные свойства параметров b и шумов $\boldsymbol{\xi} = \{\xi[1]. \dots, \xi[N]\}$ будем характеризовать совместной функцией ПРВ $W(a,\boldsymbol{\xi})$, нормированной на множестве

$$\mathcal{W} = \mathcal{B} \otimes \mathcal{Q}, \quad \mathcal{Q} = \bigotimes_{n=1}^{N} \Xi_n. \tag{31}$$

Воспользуемся моделью зависимости (1) и аддитивной моделью измерений с переменной дисперсией:

$$y[n]) = \varphi(f[n], a), \quad v[n] = y[n] + \sigma(n,b)\, \xi[n]. \tag{32}$$

Для восстановления функций ПРВ $P(a), W(a,\boldsymbol{\xi})$ также, как в п. 1, воспользуемся данными $y^r[1], \dots, y^r[N]$ и алгоритмом РМО [1], адаптированным к данной задаче:

$$\mathcal{H}[P(a), W(a,\boldsymbol{\xi})] = -\int_{\mathcal{A}} P(a) \ln P(a)\, da - \int_{\mathcal{W}} W(b,\boldsymbol{\xi}) \ln W(b,\boldsymbol{\xi}) db\, d\boldsymbol{\xi} \Rightarrow \max, \tag{33}$$

при ограничениях:

$$\int_{\mathcal{A}} P(a)\, da = 1, \quad \int_{\mathcal{W}} W(b,\boldsymbol{\xi})\, db\, d\boldsymbol{\xi} = 1, \tag{34}$$

$$\int_{\mathcal{A}} P(a)\, \varphi(f[n], a)\, da + \int_{\mathcal{W}} W(b,\boldsymbol{\xi})\, \sigma(n,b)\, \xi[n]\, da\, d\boldsymbol{\xi} = y^r[n], \quad n = \overline{1,N} \tag{35}$$

Решения этой задачи имеют следующий вид оптимальных функций ПРВ:

- параметров зависимости в (32)

$$P^*(a) = \frac{\exp\left(-\sum_{n=1}^{N} \lambda_n \varphi(f[n], a)\right)}{\mathbb{P}(\boldsymbol{\lambda})}, \quad \mathbb{P}(\boldsymbol{\lambda}) = \int_{\mathcal{A}} \exp\left(-\sum_{n=1}^{N} \lambda_n \varphi(f[n], a)\right) da; \tag{36}$$

- параметров модели дисперсии и измерительного шума

$$W^*(b, \boldsymbol{\xi}) = \frac{\exp\left(-\sum_{n=1}^{N} \lambda_n \sigma(n, b)\, \xi[n]\right)}{\mathbb{W}(\boldsymbol{\lambda})}, \tag{37}$$
$$\mathbb{W}(\boldsymbol{\lambda}) = \int_{\mathcal{W}} \exp(-\sum_{n=1}^{N} \lambda_n \sigma(n, b)\, \xi[n])\, db\, d\boldsymbol{\xi}.$$

Множители Лагранжа $\boldsymbol{\lambda}$ определяются решением балансовых уравнений:

$$\mathbb{P}^{-1}(\boldsymbol{\lambda}) \int_{\mathcal{A}} \exp(-\sum_{n=1}^{N} \lambda_n \varphi(f[n], a))\, \varphi(f[n], a)\, da +$$
$$+\mathbb{W}^{-1}(\boldsymbol{\lambda}) \int_{\mathcal{W}} \exp(-\sum_{n=1}^{N} \lambda_n \sigma(n, b)\, \xi[n])\, \sigma(n, b)\, \xi[n]\, d\xi[n] = y^r[n],$$
$$n = \overline{1, N}. \tag{38}$$

Используя последние равенства и модель дисперсии, можно вычислить средние и дисперсии параметров модели и измерительного шума. Имеем средние векторы параметров:

$$b^n = \int_{\mathcal{W}} W^*(b, \boldsymbol{\xi})\, b\, db\, d\boldsymbol{\xi}[n], \tag{39}$$

и измерительного шума:

$$\bar{\boldsymbol{\xi}}^n = \int_{\mathcal{W}} W^*(b, \boldsymbol{\xi})\, \boldsymbol{\xi}[n]\, d\boldsymbol{\xi}[n]\, db. \tag{40}$$

Дисперсии указанных векторов имеют вид:

$$D(b)_n = \int_{\mathcal{W}} W^*(b, \boldsymbol{\xi})\, (b - b_n)^2\, db\, d\boldsymbol{\xi}[n],$$
$$D(\boldsymbol{\xi})_n = \int_{\mathcal{W}} W^*(b, \boldsymbol{\xi})\, (\boldsymbol{\xi}[n] - \bar{\boldsymbol{\xi}}_n)^2\, d\boldsymbol{\xi}[n]\, db. \tag{41}$$

Рассмотрим последние равенства и модель дисперсии (32). Введем следующие обозначения:

$$W^*_{n,k}(\lambda_n, b_k, \xi[n]) = \frac{\exp\left(-\lambda_n\, n^k\, b_k\, \xi[n]\right)}{\mathcal{W}_{n,k}(\lambda_n)}, \tag{42}$$
$$\mathcal{W}_{n,k}(\lambda_n) = \int_{\mathcal{B}_k \otimes \mathcal{Q}_n} \exp\left(-\lambda_n\, n^k\, b_k\, \xi[n]\right) db_k\, d\xi[n], \quad n = \overline{1, N}, k = \overline{0, s}.$$

Тогда функции ПРВ (37) можно представить в виде:

$$W^*(b, \boldsymbol{\xi}) = \prod_{n=1}^{N} \prod_{k=0}^{s} W^*_{n,k}(\lambda_n, b_k, \xi[n]). \tag{43}$$

Из (39) следует, что частные функции ПРВ в (43) экспоненциального класса с билинейными аргументами, т.е. средние и дисперсии соответствующих компонент в (39-43) вычисляются аналитически.

Пример. Рассмотрим нелинейную зависимость

$$y[n] = \varphi(f[n], a_1, a_2) = a_1 f[n] + a_2 f^2[n] \tag{44}$$

и гетероскедастический шум

$$v[n] = y[n] + \sigma(n, b_0, b_1)\, \xi[n] = y[n] + (b_0 + b_1\, n)\, \xi[n]. \tag{45}$$

Параметры модели зависимости случайные, независимые, интервальные:

$$a_1 \in [a_1^-, a_1^+], \; a_2 \in [a_2^-, a_2^+]. \tag{46}$$

Параметры модели дисперсии также случайные, независимые и интервальные:

$$b_0 \in [0, b_0^+], \; b_1 \in [0, b_2^+]. \tag{47}$$

Допустим, что имеются данные

$$f[1], \dots, f[N]; \quad y^r[1], \dots, y^r[N], \tag{48}$$

на интервале $[1, N]$, где $N = 5$.

Согласно (36) функция ПРВ параметров зависимости имеет вид:

$$P^*(a_1, a_2 \mid \boldsymbol{\lambda}) = p_1^*(a_1 \mid \boldsymbol{\lambda})\, p_2^*(a_2 \mid \boldsymbol{\lambda}), \tag{49}$$

где:

$$p_1^*(a_1 \mid \boldsymbol{\lambda}) = w(\boldsymbol{\lambda}) \exp(-a_1\, c_1(\boldsymbol{\lambda})), \quad c_1(\boldsymbol{\lambda}) = \sum_{n=1}^{5} \lambda_n\, f[n],$$
$$p_2^*(a_2 \mid \boldsymbol{\lambda}) = w(\boldsymbol{\lambda}) \exp(-a_2\, c_2(\boldsymbol{\lambda})), \quad c_2(\boldsymbol{\lambda}) = \sum_{n=1}^{5} \lambda_n\, f^2[n]. \tag{50}$$

$$w(\boldsymbol{\lambda}) = \sqrt{\mathbb{P}(\boldsymbol{\lambda})}. \tag{51}$$

Перейдем к рассмотрению гетероскедастического измерительного шума в (32), где

$$\sigma(n, b) = b_0 + n\, b_1. \tag{52}$$

Коэфициенты в этом равенстве случайные, независимые интервальные:

$$b_0 \in [0, b_0^+], \quad b_1 \in [0, b_1^+]. \tag{53}$$

Шум $\xi[n]$ - случайный, независиммый, интервальный:

$$\xi[n] \in [\xi^-[n], \xi^+[n]]. \tag{54}$$

Совместная функция ПРВ $W(b_0, b_1; \boldsymbol{\xi})$ определена на множестве

$$\mathcal{W} = \mathcal{W}_0 \otimes \mathcal{W}_1, \tag{55}$$

$$\mathcal{W}_0 = \mathcal{B}_0 \otimes \left(\bigotimes_{n=1}^{5} \Xi_n \right), \quad \mathcal{W}_1 = \mathcal{B}_1 \otimes \left(\bigotimes_{n=1}^{5} \Xi_n \right).$$

Согласно (42) имеем:

$$W_{n,0}(\lambda_n, b_0, \xi[n]) = \mathcal{W}_{n,0}^{-1}(\lambda_n) \exp(-b_0\, \lambda_n\, \xi[n]),$$

$$W_{n,1}(\lambda_n, b_1, \xi[n]) = \mathcal{W}_{n,1}^{-1}(\lambda_n) \exp(-b_1\, \lambda_n\, n\, \xi[n]), \; n = \overline{1,5}. \tag{56}$$

В этих выражениях

$$\mathcal{W}_{n,0}(\lambda_n) = \int_{\mathcal{B}_0 \otimes \Xi_n} \exp(-b_0\, \lambda_n\, \xi[n]) db_0 d\xi[n],$$

$$\mathcal{W}_{n,1}(\lambda_n) = \int_{\mathcal{B}_1 \otimes \Xi_n} \exp(-b_1\, n\, \lambda_n\, \xi[n]) db_1 d\xi[n]. \tag{57}$$

Функция ПРВ

$$W(b_0, b_1, \boldsymbol{\xi} \mid \boldsymbol{\lambda}) = \prod_{n=1}^{5} W_{n,0}(\lambda_n, b_0, \xi[n])\, W_{n,1}(\lambda_n, b_1, \xi[n]). \tag{58}$$

Множители Лагранжа, входящие так же в (49 - 51), определяются решением балансовых уравнений.

Функция ПРВ (58) позволяет вычислить средние параметров модели дисперсии (52):

$$\bar{b}_0 = \prod_{n=1}^{n} \int_{\Xi_n} \left(\int_{\mathcal{B}_0} b_0\, W_{n,0}(\lambda_n, b_0, \xi[n])\, db_0 \right) d\xi[n],$$

$$\bar{b}_1 = \prod_{n=1}^{n} \int_{\Xi_n} \left(\int_{\mathcal{B}_1} b_1\, W_{n,0}(\lambda_n, b_1, \xi[n])\, db_1 \right) d\xi[n]. \tag{59}$$

Итак, усредненная модель наблюдений, содержащая «среднюю» модель дисперсии гетероскедастического измерительного шума приобретает следующий вид:

$$\tilde{v}[n] = y[n] + (\bar{b}_0 + n\bar{b}_1)\, \xi[n]. \tag{60}$$

5. Заключение

Исследованы модели измерительных шумов аддитивного и мультипликативного типов, а также модели зависимостей с аддитивными шумами на входе и выходе. Сформулированы теоремы об их гетероскедастической природе.

Литература

1. Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A. Entropy Randomization in Mashine Learning, 2023, CRC Press, Taylor & Francis Group.
2. Shannon C.E. Mathematical Theory of Communication. 1948, The Bell System Technical Journal, v.27, p.373-423, 623-656.
3. Jaynes E.T. Information theory and statistical Mechanics. Physical Review, 1957, v.104(4), p.620-630.
4. Jaynes E.T. Gibbs vs Boltzmann entropy. American Journal of Physics, 1965, v.33, p.391-398.
5. Rosenkrantz R.D., Jaynes E.T. Paper on Probability, Statistics,and Statistical Physics. Kluwer Academic Pablishers, 1989.
6. Popkov Yu.S. Qualitative Properties of Random Maximum Entropy Estimates of Probability Density Functions. Mathematics, 2021, 9, 548, doi.org/10.3390/math9050548.
7. Bollerslev T., Engle R.F., Nelson D.B. ARCH Models. In: Engle R.F. and McFadden D.C.,eds. Handbook of

Econometrics, 1994, Elsevier Science, Amsterdam, p.2961-3038.
8. Cai T.T., Wang L. Adaptive variance function estimation in heteroscedastic nonparametric regression. Ann. Stat., 2008, v. 36(5), p. 2025-2054, doi:10.1214/07-AOS509.
9. Орешко Н.И. Восстановление закона изменения гетероскедастического шума при траекторных измерениях на основе вейвлет-технологий. Известия СПбГЭТУ "ЛЭТИ", 2013, № 9, с. 16-21.

Попков Юрий Соломонович. Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия. Главный научный сотрудник; академик РАН, доктор технических наук, профессор. Институт проблем управления Российской академии наук, Москва, Россия. Главный научный сотрудник. Область научных интересов: энтропийные методы, макросистемы, рандомизированное машинное обучение. E-mail: popkov@isa.ru

On Some Properties of Randomized Machine Learning Procedures in the Presence of Noisy Data

Y. S. Popkov

Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia

Abstract. We study various models of measuring noises in the procedures of randomized entropy estimation of probability density functions: additive and multiplicative, measuring noises at the input and output of the object’s model. The properties of entropy-optimal probability density functions are studied, it is shown that the measurement noises corresponding to them are heteroscedastic.
Keywords: entropy estimation, density functions, Lagrange multipliers, heteroscedastic noise, variation models.

DOI 10.14357/20718632230209

References

1. Popkov Yu.S., Popkov A.Yu., Dubnov Yu.A. Entropy Randomization in Mashine Learning, 2023, CRC Press, Taylor & Francis Group.
2. Shannon C.E. Mathematical Theory of Communication. 1948, The Bell System Technical Journal, v.27, p.373-423, 623-656.
3. Jaynes E.T. Information theory and statistical Mechanics. Physical Review, 1957, v.104(4), p.620-630.
4. Jaynes E.T. Gibbs vs Boltzmann entropy. American Journal of Physics, 1965, v.33, p.391-398.
5. Rosenkrantz R.D., Jaynes E.T. Paper on Probability, Statistics,and Statistical Physics. Kluwer Academic Pablishers, 1989.
6. Popkov Yu.S. Qualitative Properties of Random Maximum Entropy Estimates of Probability Density Functions. Mathematics, 2021, 9, 548, doi.org/10.3390/math9050548.
7. Bollerslev T., Engle R.F., Nelson D.B. ARCH Models. In: Engle R.F. and McFadden D.C.,eds. Handbook of Econometrics, 1994, Elsevier Science, Amsterdam, p.2961-3038.
8. Cai T.T., Wang L. Adaptive variance function estimation in heteroscedastic nonparametric regression. Ann. Stat., 2008, v. 36(5), p. 2025-2054, doi:10.1214/07-AOS509.
9. Oreshko N.I. Vosstanovlenie zakona heteroskedaticheskogo shuma pri traektornikh izmereniah na osnove veivlet-tehnologiy // Izvestia SPbLETU “LETI”, 2013, No. 9, pp. 16-21.

Popkov Y. S.. Federal Reearch Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia, chief research scientist.; Member of RAS, Doctor of Science in Engineering, professor; Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia, chief research scientist. Scientific area: entropy methods, macrosystems, randomized machine learning. E-mail: popkov@isa.ru