Семинар «Прикладная статистика» (ИМ СО РАН)

4 апреля 2023 года

Экономичные компьютерные функциональные алгоритмы приближения вероятностных плотностей по заданной выборке

ИВМиМГ СО РАН

Войтишек Антон Вацлавович

(ИВМиМГ СО РАН, г. Новосибирск)

ПЛАН ДОКЛАДА

1. О группе исследователей по методам Монте-Карло в ИВМиМГ СО РАН (Новосибирск) и о соавторах представляемых результатов исследований: Е. В. Шкарупа, В. В. Милосердове, Е. Г. Каблуковой, Т. Е. Булгаковой, Н. Х. Шлымбетове.

2. Об относительной простоте и специфике основных вычислительных конструкций метода Монте-Карло. Функциональные алгоритмы метода Монте-Карло.

3. О задаче «быстрого» приближения вероятностной плотности по заданной выборке. Насколько она актуальна?!

4. Использование классических компьютерных сеточных приближений. О выборе аппроксимационного базиса.

5. Компьютерное приближение значений плотности в узлах сетки: использование ядерных и проекционных непараметрических оценок. Важный частный случай: многомерный аналог полигона частот. Есть ли другие конструктивные подходы?!

6. Применение теории условной оптимизации вычислительных функциональных алгоритмов. Преимущества многомерного аналога полигона частот с точки зрения использования этой теории.

Основная схема метода Монте-Карло (см., например, [1]):

$$I = \mathbf{E}\zeta \approx Z_n = \frac{\zeta_1 + \cdots + \zeta_n}{n}.$$

Погрешность (доверительный интервал) $\mathbf{P}\left\{|I - Z_n| \le H_\varepsilon \frac{\sqrt{\mathbf{D}\zeta}}{\sqrt{n}}\right\} > 1 - \varepsilon.$

В данном докладе существенно используется методология оптимизации **функциональных алгоритмов метода Монте-Карло** (см. прежде всего [2, 3]).

Эти алгоритмы используются для компьютерного приближения решения интегрального уравнения Фредгольма второго рода

$$\varphi(x) = \int k(x', x)\varphi(x')\,dx' + f(x)$$

с использованием различных вариантов т. н. **оценки по столкновениям**

$$(\varphi, h) = \int \varphi(x)h(x)dx = \mathbf{E}\sum_{m=0}^{N} Q^{(m)} h\big(\xi^{(m)}\big),$$

где $\xi^{(0)}, \xi^{(1)}, \ldots, \xi^{(N)}$ – т. н. **«прикладная» (обрывающаяся) цепь Маркова.**

[1] *Войтишек А. В.* Лекции по численным методам Монте-Карло. – Новосибирск: ИПЦ НГУ, 2018 Войтишек-лекции-по-ММК.pdf (nsu.ru)

[2] *Войтишек А. В.* Функциональные оценки метода Монте-Карло. – Новосибирск: НГУ, 2007 voytishekav-functional-estimations-of-monte-carlo-methods.pdf (nsu.ru)

[3] *Шкарупа Е. В.* Сходимость и оптимизация численных дискретно-стохастических процедур. – Новосибирск, 2000 (Диссертация на соискание ученой степени кандидата физико-математических наук / НГУ).

При обработке больших данных (в частности, при применении технологий машинного обучения) актуальной может оказаться следующая задача.

ЗАДАЧА 1 **[4, 5]**. *По заданной выборке* $\{\xi_1, \ldots, \xi_n\}$ *построить численное (компьютерное) функциональное приближение неизвестной плотности* $f_{\xi}(\boldsymbol{x})$ *на компактной области распределения* $X \subset \mathbb{R}^d$ *случайной величины (вектора)* $\boldsymbol{\xi} \in X$ *с заданным уровнем погрешности* $L > 0$ *и с наименьшими вычислительными затратами* S.

Отличие от функциональных алгоритмов: значения $\{\xi_1, \ldots, \xi_n\}$ не моделируются, а заданы.

[4] *Булгакова Т. Е., Войтишек А. В.* Условная оптимизация функционального вычислительного ядерного алгоритма приближения вероятностной плотности по заданной выборке // Журнал вычислительной математики и математической физики. – 2021. – Т. 61, № 9. – С. 29–44.

[5] *Булгакова Т. Е.* Оптимизация функциональных вычислительных статистических оценок и алгоритмов. – Новосибирск, 2020 (Диссертация на соискание ученой степени кандидата физико-математических наук / ИВМиМГ СО РАН).

При компьютерном приближении плотности $f_\xi(\boldsymbol{x})$ будем следовать классическим численным схемам приближения функций **[6]** и рассматривать приближения вида

$$f_\xi(\boldsymbol{x}) \approx L^{(M)} f_\xi(\boldsymbol{x}) = \sum_{i=1}^{M} w^{(i)}\left[f_\xi(\boldsymbol{y}_1), \dots, f_\xi(\boldsymbol{y}_M)\right] \chi^{(i)}(\boldsymbol{x});$$

$Y^{(M)} = \{\boldsymbol{y}_1, \dots, \boldsymbol{y}_M\}$ – аппроксимационная сетка (как правило, регулярная),

$\Xi^{(M)} = \{\chi^{(i)}(\boldsymbol{x}), \dots, \chi^{(i)}(\boldsymbol{x})\}$ – аппроксимационный базис,

$W^{(M)} = \{w^{(i)}\left[f_\xi(\boldsymbol{y}_1), \dots, f_\xi(\boldsymbol{y}_M)\right]; i = 1, \dots, M\}$ – коэффициенты, являющиеся композициями (как правило, линейными комбинациями) значений $f_\xi(\boldsymbol{y}_1), \dots, f_\xi(\boldsymbol{y}_M)$ функции $f_\xi(\boldsymbol{x})$ в узлах сетки $Y^{(M)}$.

Нами с Н. Х. Шлымбетовым, с использованием соображений из статьи **[7]**, проводятся исследования аппроксимационных базисов, подходящих для решения Задачи 1.

[6] *Бахвалов Н. С.* Численные методы. – М.: Наука, 1975.

[7] *Voytishek A. V., Kablukova E. G.* Using the approximation functional bases in Monte Carlo methods // Russian Journal of Numerical Analysis and Mathematical Modelling. – 2003. – V. 18, № 6. – P. 521–542.

Некоторые важные принципы выбора аппроксимационного базиса.

А). Использование мультипликативных конструкций, в частности, конструкций из [8] (базисных функций аппроксимации Стренга-Фикса)

$$\chi^{(i)}(\boldsymbol{x}) = B^{(r)}\left[\frac{x^{(1)}}{h} - j_i^{(1)}\right] \times \dots \times B^{(r)}\left[\frac{x^{(d)}}{h} - j_i^{(d)}\right]; \qquad (*)$$

$\boldsymbol{x} = \left(x^{(1)}, \dots, x^{(d)}\right)$; $\boldsymbol{y}_i = \left(j_i^{(1)}h, \dots, j_i^{(d)}h\right)$, где $j_i^{(k)}$ – целые числа; $B^{(r)}(w)$ - это B-сплайн r-го порядка; это позволяет проще переходить от одномерного к многомерным случаям.

Б). Нужны приемлемые аппроксимационные свойства. Например, для базиса (*)

$$\left\|f_\xi - L^{(M)} f_\xi\right\|_{\mathbb{C}(X)} \le H^{(r)} h^{r+1} \left\|f_\xi\right\|_{\mathbb{C}^{(r+1)}(X)}.$$

В). Использовать базисы $\Xi^{(M)} = \left\{\chi^{(i)}(\boldsymbol{x}), \dots, \chi^{(i)}(\boldsymbol{x})\right\}$ с небольшой константой Лебега $H^{(Leb)} = \sum_{i=1}^{M}\left|\chi^{(i)}(\boldsymbol{x})\right|$. Для базиса (*) имеем $H^{(Leb)} = 1$.

Г). Использовать базисы $\Xi^{(M)}$, которым соответствуют «компактные» коэффициенты; лучше всего, чтобы выполнялось $w^{(i)}\left[f_\xi(\boldsymbol{y}_1), \dots, f_\xi(\boldsymbol{y}_M)\right] = f_\xi(\boldsymbol{y}_i); i = 1, \dots, M$.

В), Г) - это соображения устойчивости.

[8] Марчук Г. И., Агошков В. И. Введение в проекционно-сеточные методы. – М.: Наука, 1981.

Проведенные нами исследования (изучались базисы Лагранжа и Бернштейна, тригонометрические базисы, базис Стренга-Фикса для $r = 1, 3$) и с учетом замечаний из работ [6, 7, 9], позволяют сделать вывод о целесообразности использования **мульти-линейного приближения**

$$f_\xi(\boldsymbol{x}) \approx L^{(M)} f_\xi(\boldsymbol{x}) = \sum_{i=1}^{M} f_\xi(\boldsymbol{y}_i)\, \chi^{(i)}(\boldsymbol{x})\,;$$

$$\chi^{(i)}(\boldsymbol{x}) = B^{(1)}\left[\frac{x^{(1)}}{h} - j_i^{(1)}\right] \times \dots \times B^{(1)}\left[\frac{x^{(d)}}{h} - j_i^{(d)}\right]; \boldsymbol{x} = \left(x^{(1)}, \dots, x^{(d)}\right);$$

$$B^{(1)}(u) = \begin{cases} u + 1 \text{ при } -1 \le u \le 0; \\ -u + 1 \text{ при } 0 \le u \le 1; \\ 0 \quad \text{иначе}; \end{cases} \quad \boldsymbol{y}_i = \left(j_i^{(1)} h, \dots, j_i^{(d)} h\right).$$

При решении задачи 1 нужно учесть, что, в отличие от постановок теории численного приближения функций, значения (коэффициенты) $f_\xi(\boldsymbol{y}_i)$ не заданы; **их требуется приближать по заданной выборке** $\{\xi_1, \dots, \xi_n\}$.

[9] *Милосердов В. В.* Дискретно-стохастические численные алгоритмы со сплайн-восполнениями. – Новосибирск, 2006 (Диссертация на соискание ученой степени кандидата физико-математических наук / НГУ).

Имеется два основных способа приближения значения плотности $f_\xi(\boldsymbol{x})$ в заданной точке $\boldsymbol{x} = \widehat{\boldsymbol{y}}$ по заданной выборке $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$.

Первый способ – использование **ядерной (точечной) статистической оценки плотности**

$$f_\xi(\widehat{\boldsymbol{y}}) \approx \tilde{f}_\xi^{(\widehat{\boldsymbol{y}};ker)}(n) = \frac{1}{n}\sum_{j=1}^{n} \kappa^{(\widehat{\boldsymbol{y}})}\left(\boldsymbol{\xi}_j\right);$$

здесь $\kappa^{(\boldsymbol{y})}(\boldsymbol{z})$ – специально выбираемая *ядерная функция*, такая, что $f_\xi(\widehat{\boldsymbol{y}}) \approx \int_X f_\xi(\boldsymbol{z})\kappa^{(\widehat{\boldsymbol{y}})}(\boldsymbol{z})\,d\boldsymbol{z} = \mathbf{E}\kappa^{(\widehat{\boldsymbol{y}})}(\boldsymbol{\xi})$. Этот способ дает следующий алгоритм.

АЛГОРИТМ 1 (**[4, 5]**, ***вычислительный функциональный ядерный алгоритм приближения вероятностной плотности***). *Приближаем значения* $f_\xi(\boldsymbol{y}_i)$ *по формуле* $f_\xi(\boldsymbol{y}_i) \approx \tilde{f}_\xi^{(\boldsymbol{y}_i;ker)}(n) = \frac{1}{n}\sum_{j=1}^{n} \kappa^{(\boldsymbol{y}_i)}\left(\boldsymbol{\xi}_j\right); i = 1, \dots, M$ *(здесь* $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$ *– заданная выборка) и строим окончательную аппроксимацию плотности* $f_\xi(\boldsymbol{x})$*:*

$$f_\xi(\boldsymbol{x}) \approx L^{(M,n)}\tilde{f}_\xi^{(ker)}(\boldsymbol{x}) = \sum_{i=1}^{M} \tilde{f}_\xi^{(\boldsymbol{y}_i;ker)}(n)\,\chi^{(i)}(\boldsymbol{x}).$$

Второй способ приближения значения плотности $f_{\xi}(\boldsymbol{x})$ в заданной точке $\boldsymbol{x} = \widehat{\boldsymbol{y}}$ по заданной выборке $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$ – использование **проекционной (точечной) статистической оценки плотности**

$$f_{\xi}(\widehat{\boldsymbol{y}}) \approx \tilde{f}_{\xi}^{(\widehat{\boldsymbol{y}};pr)}(n, K) = \sum_{k=1}^{K} \left[\frac{1}{n}\sum_{j=1}^{n} \psi^{(k)}(\boldsymbol{\xi}_j)\right] \psi^{(k)}(\widehat{\boldsymbol{y}});$$

здесь $\boldsymbol{\Psi}^{(K)} = \{\psi^{(1)}(\boldsymbol{y}), \dots, \psi^{(K)}(\boldsymbol{y})\}$ – подмножество бесконечной ортонормированной системы функций $\boldsymbol{\Psi}^{(\infty)} = \{\psi^{(1)}(\boldsymbol{y}), \psi^{(2)}(\boldsymbol{y}), \dots\}$ такой, что $\int_X \psi^{(i)}(\boldsymbol{y})\,\psi^{(j)}(\boldsymbol{y})\,d\boldsymbol{y} = 1$ при $i = j$; $\int_X \psi^{(i)}(\boldsymbol{y})\,\psi^{(j)}(\boldsymbol{y})\,d\boldsymbol{y} = 0$ при $i \neq j$ и $f_{\xi}(\boldsymbol{x}) = \sum_{k=1}^{\infty} c^{(k)}\psi^{(k)}(\boldsymbol{x}) = \sum_{k=1}^{\infty} \mathbf{E}\psi^{(k)}(\boldsymbol{\xi})\,\psi^{(k)}(\boldsymbol{x})$. Получаем алгоритм.

АЛГОРИТМ 2 (**[10]**, ***вычислительный функциональный проекционный алгоритм приближения вероятностной плотности***). *Приближаем значения* $f_{\xi}(\boldsymbol{y}_i)$ *по формуле* $f_{\xi}(\boldsymbol{y}_i) \approx \tilde{f}_{\xi}^{(\boldsymbol{y}_i;pr)}(n, K) = \sum_{k=1}^{K} \left[\frac{1}{n}\sum_{j=1}^{n} \psi^{(k)}(\boldsymbol{\xi}_j)\right] \psi^{(k)}(\boldsymbol{y}_i)$; $i = 1, \dots, M$ *и строим окончательную аппроксимацию плотности* $f_{\xi}(\boldsymbol{x})$*:*

$$f_{\xi}(\boldsymbol{x}) \approx L^{(M,n,K)}\tilde{f}_{\xi}^{(pr)}(\boldsymbol{x}) = \sum_{i=1}^{M} \tilde{f}_{\xi}^{(\boldsymbol{y}_i;pr)}(n, K)\,\chi^{(i)}(\boldsymbol{x}).$$

[10] *Войтишек А. В., Шлымбетов Н. Х.* Экономичные алгоритмы приближения вероятностной плотности по заданной выборке // Материалы XXI Международной конференции имени А. Ф. Терпугова «Информационные технологии и математической моделирование» (Карши, Узбекистан, 25–29 октября 2022 года). – Томск: ТГУ, 2023 *(в печати)*.

Сформулируем следующее методически важное замечание [10].

Если взять одновременно ядерную функцию $\kappa^{(x)}(\boldsymbol{z})$ в виде

$$\kappa^{(x)}(\boldsymbol{z}) = \frac{\boldsymbol{I}^{\left(\Delta^{(x)}\right)}(\boldsymbol{z})}{h^d};\ \ \boldsymbol{I}^{(A)}(\boldsymbol{z}) = \begin{cases} 1 \text{ при } \boldsymbol{z} \in A, \\ 0 \text{ иначе;} \end{cases}$$

$$\Delta^{(\boldsymbol{x})} = \left\{\boldsymbol{z} = \left(z^{(1)}, \ldots, z^{(d)}\right): x^{(s)} - \frac{h}{2} \le z^{(s)} < x^{(s)} + \frac{h}{2};\ s = 1, \ldots, d;\ \boldsymbol{x} = \left(x^{(1)}, \ldots, x^{(d)}\right)\right\}$$

(здесь $\boldsymbol{I}^{(A)}(\boldsymbol{z})$ – индикатор множества A) и функции $\boldsymbol{\Psi}^{(\infty)}$ вида

$$\psi^{(i)}(\boldsymbol{z}) = \frac{\boldsymbol{I}^{\left(\Delta^{(y_i)}\right)}(\boldsymbol{z})}{h^d},$$

то получается следующая хорошо изученная вычислительная схема.

АЛГОРИТМ 3 ([4, 5], ***многомерный аналог полигона частот***). *Приближаем значения $f_\xi(\boldsymbol{y}_i)$ по формулам $f_\xi(\boldsymbol{y}_i) \approx \tilde{f}_\xi^{(y_i;pol)}(n) = \frac{1}{nh^d}\sum_{j=1}^n \boldsymbol{I}^{\left(\Delta^{(y_i)}\right)}(\boldsymbol{\xi}_j)$; $i = 1, \ldots, M$ (здесь $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n\}$ – заданная выборка) и строим окончательную аппроксимацию плотности $f_\xi(\boldsymbol{x})$:*

$$f_\xi(\boldsymbol{x}) \approx L^{(M,n)}\tilde{f}_\xi^{(pol)}(\boldsymbol{x}) = \sum_{i=1}^M \tilde{f}_\xi^{(y_i;pol)}(n)\, \chi^{(i)}(\boldsymbol{x}).$$

Соавтор данных исследований, аспирант НГУ **Н. Х. Шлымбетов**, в рамках подготовки кандидатской диссертации

«Выбор аппроксимационных базисов и специальных систем функций в вычислительных алгоритмах приближения вероятностных плотностей», будет проводить сравнительное изучение сформулированных вычислительных функциональных ядерных и проекционных алгоритмов приближения вероятностной плотности по заданной выборке (алгоритмы 1 и 2), отличных от метода полигона частот (алгоритма 3), а конкретнее, изучение возможностей построения **теории условной оптимизации** (см. прежде всего **[2–5])** для алгоритмов 1 и 2 в сравнении с хорошо изученным в **[4, 5]** алгоритмом 3.

Предположительно, основной вывод диссертационной работы будет состоять в том, что **использование вариантов алгоритмов 1 и 2, отличных от многомерного аналога полигона частот (алгоритма 3), приводит к существенным проблемам** в разработке подходов по согласованному выбору нужного объема n выборки $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$, числа M узлов сетки $\{\boldsymbol{y}_i\}$ для достижения заданного уровня погрешности $L > 0$ с наименьшими затратами S (для проекционного алгоритма 2 требуются также соображения по разумному выбору числа K используемых ортонормированных функций $\boldsymbol{\Psi}^{(K)} = \{\psi^{(1)}(\boldsymbol{y}), \dots, \psi^{(K)}(\boldsymbol{y})\}$).

Упомянутые выше проблемы возникают в рамках следующей **общей схемы теории условной оптимизации** [2–5].

Ставится задача согласованного выбора параметров M и n функционального алгоритма, обеспечивающего заданный уровень L погрешности $\delta^{(\mathbb{B})}(M,n)$ (для подходящего нормированного функционального пространства $\mathbb{B}(X)$) приближения

$$L^{(M,n)}\tilde{f}_\xi(\boldsymbol{x}) = L^{(M,n)}\tilde{f}_\xi^{(ker)}(\boldsymbol{x}) \vee L^{(M,n,K)}\tilde{f}_\xi^{(pr)}(\boldsymbol{x}) \vee L^{(M,n)}\tilde{f}_\xi^{(pol)}(\boldsymbol{x})$$

при минимальных вычислительных затратах $S(M,n)$.

МЕТОД 1 **[2–5]**. *Строится верхняя граница $UP^{(\mathbb{B})}(M,n)$ погрешности $\delta^{(\mathbb{B})}(M,n)$, зависящая от параметров M и n:*

$$\delta^{(\mathbb{B})}(M,n) = \left\| f_\xi - L^{(M,n)}\tilde{f}_\xi \right\|_{\mathbb{B}(X)} \le UP^{(\mathbb{B})}(M,n).$$

Эта функция двух переменных приравнивается величине L. Из уравнения вида $UP^{(\mathbb{B})}(M,n) = L$ один из параметров (например, n) выражается через другой: $n = Y^{(L)}(M)$.

Это соотношение подставляется в выражение для затрат $S(M,n)$ (которое тоже зависит от параметров M и n). В результате получается функция $\tilde{S}^{(\mathbb{B},L)}(M)$ одного переменного M, которая исследуется на минимум с помощью известных приемов математического или численного анализа.

Найденные значения $M_{min}^{(\mathbb{B})}(L) = M_{opt}^{(\mathbb{B})}(L)$, $n_{opt}^{(\mathbb{B})}(L) = Y^{(L)}\left[M_{opt}^{(\mathbb{B})}\right]$ объявляются ***условно-оптимальными параметрами*** *соответствующего вычислительного алгоритма.*

ЗАМЕЧАНИЕ. При оптимизации алгоритмов 1 и 2, кроме выбора параметров M и n, важным является оптимальный, согласованный с параметрами $M_{opt}^{(\mathbb{B})}(L)$ и $n_{opt}^{(\mathbb{B})}(L)$, выбор вспомогательных функций $\kappa^{(\boldsymbol{x})}(z)$ и $\boldsymbol{\Psi}^{(K)} = \{\psi^{(1)}(y), \dots, \psi^{(K)}(y)\}$, соответственно. Это приводит к рассмотрению дополнительных параметров (например, параметра K) и к определенной смене методики оптимизации (см. далее метод 2).

При изучении погрешности $\delta^{(\mathbb{B})}(M,n)$ необходимо выбрать как соответствующее нормированное функциональное пространство $\mathbb{B}(X)$, так и вероятностный смысл выполнения неравенства вида $\delta^{(\mathbb{B})}(M,n) \le UP^{(\mathbb{B})}(M,n)$ (ведь $\delta^{(\mathbb{B})}(M,n)$ является случайной величиной). Следуя канонам классического численного анализа, в качестве пространств $\mathbb{B}(X)$ будем рассматривать $\mathbb{L}_2(X)$ и $\mathbb{C}(X)$.

Для хорошо разработанного $\mathbb{L}_2$ -**подхода** (см., например, **[2–5]**) выбирается сходимость в среднем погрешности

$$\delta^{(\mathbb{L}_2)}(M,n) = \left\| f_\xi - L^{(M,n)}\tilde{f}_\xi \right\|_{\mathbb{L}_2(X)} = \left(\int_X \left[f_\xi(\boldsymbol{x}) - L^{(M,n)}\tilde{f}_\xi(\boldsymbol{x}) \right]^2 d\boldsymbol{x} \right)^{1/2}$$

к нулю при $M, n \to \infty$ и строятся оценки сверху $UP^{(\mathbb{L}_2)}(M,n)$ такие, что

$$\left[\mathbf{E}\delta^{(\mathbb{L}_2)}(M,n) \right]^2 \le UP^{(\mathbb{L}_2)}(M,n).$$

Для $\mathbb{C}$-**подхода** (см., например, [2–5]) величина

$$\delta^{(\mathbb{C})}(M,n) = \left\| f_\xi - L^{(M,n)} \tilde{f}_\xi \right\|_{\mathbb{C}(X)} = \sup_{\boldsymbol{x} \in X} \left| f_\xi(\boldsymbol{x}) - L^{(M,n)} \tilde{f}_\xi(\boldsymbol{x}) \right|$$

ограничивается сверху по вероятности:

$$\mathbf{P}\left[\delta^{(\mathbb{C})}(M,n) \le UP^{(\mathbb{C})}(M,n)\right] > 1 - \varepsilon$$

для некоторого достаточно малого $\varepsilon > 0$.

Для $\mathbb{L}_2$-подхода и $\mathbb{C}$-подхода верхнюю границу погрешности $\delta^{(\mathbb{B})}(M,n)$ для $\mathbb{B}(X) = \mathbb{C}(X) \vee \mathbb{L}_2(X)$ записывают в виде суммы трех компонент:

– **компоненты аппроксимации** $\delta^{(\mathbb{B})}_{appr}(M)$ (в работах [2–5] она названа *детерминированной компонентой*, что приводит к некоторой «путанице»),

– **стохастической компоненты** $\delta^{(\mathbb{B})}_{stoch}(M,n)$ и

– **компоненты смещения** $\delta^{(\mathbb{B})}_{bias}(M)$.

Для $\mathbb{C}$-подхода имеем

$$\delta^{(\mathbb{C})}(M,n) = \sup_{\boldsymbol{x}\in X}\left|f_{\boldsymbol{\xi}}(\boldsymbol{x}) - L^{(M,n)}\tilde{f}_{\boldsymbol{\xi}}(\boldsymbol{x})\right| \le \delta_{appr}^{(\mathbb{C})}(M) + \delta_{stoch}^{(\mathbb{C})}(M,n) + \delta_{bias}^{(\mathbb{C})}(M),$$

где $\delta_{appr}^{(\mathbb{C})}(M) = \left\|f_{\boldsymbol{\xi}} - L^{(M)}f_{\boldsymbol{\xi}}\right\|_{\mathbb{C}(X)}$, $\delta_{stoch}^{(\mathbb{C})}(M,n) = \left\|L^{(M)}\bar{f}_{\boldsymbol{\xi}} - L^{(M,n)}\tilde{f}_{\boldsymbol{\xi}}\right\|_{\mathbb{C}(X)}$, $\delta_{bias}^{(\mathbb{C})}(M) = \left\|L^{(M)}f_{\boldsymbol{\xi}} - L^{(M)}\bar{f}_{\boldsymbol{\xi}}\right\|_{\mathbb{C}(X)}$ (см., например, **[2–5]**).

Здесь

$$L^{(M)}\bar{f}_{\boldsymbol{\xi}}(\boldsymbol{x}) = L^{(M)}\bar{f}_{\boldsymbol{\xi}}^{(ker)}(\boldsymbol{x}) = \textstyle\sum_{i=1}^{M}\mathbf{E}\kappa^{(\boldsymbol{y}_i)}(\boldsymbol{\xi})\,\chi^{(i)}(\boldsymbol{x})$$ для алгоритма 1,

$$L^{(M)}\bar{f}_{\boldsymbol{\xi}}(\boldsymbol{x}) = L^{(M,K)}\bar{f}_{\boldsymbol{\xi}}^{(pr)}(\boldsymbol{x}) = \sum_{i=1}^{M}\left[\sum_{k=1}^{K}\mathbf{E}\psi^{(k)}(\boldsymbol{\xi})\psi^{(k)}(\boldsymbol{y}_i)\right]\chi^{(i)}(\boldsymbol{x})$$

для алгоритма 2 и

$$L^{(M)}\bar{f}_{\boldsymbol{\xi}}(\boldsymbol{x}) = L^{(M)}\bar{f}_{\boldsymbol{\xi}}^{(pol)}(\boldsymbol{x}) = \frac{1}{h^d}\sum_{i=1}^{M}\mathbf{E}\boldsymbol{I}^{\left(\Delta^{(\boldsymbol{y}_i)}\right)}(\boldsymbol{\xi})\,\chi^{(i)}(\boldsymbol{x})$$ для алгоритма 3.

Для $\mathbb{L}_2$-подхода получаем

$$\left[\mathbf{E}\delta^{(\mathbb{L}_2)}(M,n)\right]^2 = \left[\mathbf{E}\left(\int_X \left[f_\xi(\boldsymbol{x}) - L^{(M,n)}\tilde{f}_\xi(\boldsymbol{x})\right]^2 d\boldsymbol{x}\right)^{\frac{1}{2}}\right]^2 \leq$$
$$\leq 2\left[\delta_{appr}^{(\mathbb{L}_2)}(M)\right]^2 + \delta_{stoch}^{(\mathbb{L}_2)}(M,n) + 2\left[\delta_{bias}^{(\mathbb{L}_2)}(M)\right]^2,$$

где $\delta_{appr}^{(\mathbb{L}_2)}(M) = \left\|f_\xi - L^{(M)}f_\xi\right\|_{\mathbb{L}_2(X)}$, $\delta_{stoch}^{(\mathbb{L}_2)}(M,n) = \int_X \mathbf{D}L^{(M,n)}\tilde{f}_\xi(\boldsymbol{x})\,d\boldsymbol{x}$, $\delta_{bias}^{(\mathbb{L}_2)}(M) = \left\|L^{(M)}f_\xi - L^{(M)}\bar{f}_\xi\right\|_{\mathbb{L}_2(X)}$ (см., например, [2–5]).

Проведенный нами подробный сравнительный анализ построения верхних границ для компонент погрешности алгоритмов 1–3 $\delta_{appr}^{(\mathbb{B})}(M) = \delta_{appr}^{(\mathbb{B},ker)}(M) \vee \delta_{appr}^{(\mathbb{B},pr)}(M) \vee \delta_{appr}^{(\mathbb{B},pol)}(M)$, $\delta_{bias}^{(\mathbb{B})}(M) = \delta_{bias}^{(\mathbb{B},ker)}(M) \vee \delta_{bias}^{(\mathbb{B},pr)}(M,K) \vee \delta_{bias}^{(\mathbb{B},pol)}(M)$ и $\delta_{stoch}^{(\mathbb{B})}(M,n) = \delta_{stoch}^{(\mathbb{B},ker)}(M,n) \vee \delta_{stoch}^{(\mathbb{B},pr)}(M,n,K) \vee \delta_{stoch}^{(\mathbb{B},pol)}(M,n)$ для $\mathbb{B}(X) = \mathbb{L}_2(X) \vee \mathbb{C}(X)$ показывает, что оптимизационный **метод 1 реализуем только для алгоритма 3**, а для алгоритмов 1 и 2 следует применять

МЕТОД 2. *Выбор параметров* M *и* n *(а также вспомогательных функций* $\kappa^{(\boldsymbol{x})}(z)$ *и* $\boldsymbol{\Psi}^{(K)} = \left\{\psi^{(1)}(y), \dots, \psi^{(K)}(y)\right\}$*) происходит согласно соотношениям*

$$\delta_{appr}^{(\mathbb{B})}(M) = \frac{L}{3}, \qquad \delta_{bias}^{(\mathbb{B})}(M) = \frac{L}{3}, \qquad \delta_{stoch}^{(\mathbb{B})}(M,n) = \frac{L}{3}.$$

Приведем **формулы условно-оптимальных параметров, полученных в [4, 5] по методу 1 для многомерного аналога полигона частот (алгоритма 3)**

$$M_{opt}^{(\mathbb{L}_2,pol)}(L) = \left[A_1^{(\mathbb{L}_2)}\right]^{\frac{d}{4}} \times \left(\frac{d+4}{d}\right)^{\frac{d}{4}} \times L^{-d/2},$$

$$n_{opt}^{(\mathbb{L}_2,pol)}(L) = \frac{A_2^{(\mathbb{L}_2)} \times \left[A_1^{(\mathbb{L}_2)}\right]^{\frac{d}{4}} \times (d+4)^{\frac{d}{4}+1}}{4d^{\frac{d}{4}}} \times L^{-2-\frac{d}{2}},$$

$$M_{opt}^{(\mathbb{C},pol)}(L) = \left(\frac{A_1^{(\mathbb{C})} \times [(2a+1)d+4]}{(2a+1)\times d}\right)^{\frac{d}{2}} \times L^{-\frac{d}{2}},$$

$$n_{opt}^{(\mathbb{C},pol)}(L) = \frac{\left[A_1^{(\mathbb{C})}\right]^{\frac{d}{2}} \times \left[A_2^{(\mathbb{C})}\right]^2 \times [(2a+1)d+4]^{2+\frac{d}{2}}}{16[(2a+1)d]^{\frac{d}{2}}} \times$$

$$\times \left[2\ln M_{opt}^{(\mathbb{C},pol)}(L) - \ln\ln M_{opt}^{(\mathbb{C},pol)}(L) + 2A_3^{(\mathbb{C})}\right] \times L^{-2-\frac{d}{2}}$$

для некоторых положительных констант $A_1^{(\mathbb{L}_2)}, A_2^{(\mathbb{L}_2)}, A_1^{(\mathbb{C})}, A_2^{(\mathbb{C})}, A_3^{(\mathbb{C})}$ и a.

Для алгоритма 1 с произвольной ядерной функцией $\kappa^{(x)}(\boldsymbol{z})$ для выбора параметров M, n **можно предложить метод 2 с использованием**:

– для $\mathbb{C}$-подхода: соотношений

$$H_{appr}^{(\mathbb{C})} M^{-\frac{2}{d}} = \frac{L}{3}; \sup_{\boldsymbol{y} \in X} \left| \frac{1}{\text{mes } X} \int_X \left[f_\xi(\boldsymbol{z})\, \kappa^{(\boldsymbol{y})}(\boldsymbol{z}) \text{ mes } X - f_\xi(\boldsymbol{y}) \right] d\boldsymbol{z} \right| = \frac{L}{3}; \frac{H_{stoch}^{(\mathbb{C},ker)}}{\sqrt{n}} = \frac{L}{3};$$

– для $\mathbb{L}_2$-подхода: соотношений

$$2 \left[H_{appr}^{(\mathbb{L}_2)} \right]^2 M^{-\frac{4}{d}} = \frac{L^2}{3}; \qquad \frac{\text{mes } X \times \max\limits_{i=1,\ldots,M} \mathbf{D}\kappa^{(\boldsymbol{y}_i)}(\boldsymbol{\xi})}{n} = \frac{L^2}{3};$$

$$\frac{2}{\text{mes } X} \sup_{\boldsymbol{y} \in X} \left(\int_X \left[f_\xi(\boldsymbol{z})\, \kappa^{(\boldsymbol{y})}(\boldsymbol{z}) \text{ mes } X - f_\xi(\boldsymbol{y}) \right] d\boldsymbol{z} \right)^2 = \frac{L^2}{3}$$

для некоторых положительных констант $H_{appr}^{(\mathbb{C})}, H_{stoch}^{(\mathbb{C},ker)}, H_{appr}^{(\mathbb{L}_2)}$.

Для алгоритма 2 с ортонормированными функциями $\boldsymbol{\Psi}^{(K)} = \{\psi^{(1)}(\boldsymbol{y}), \dots, \psi^{(K)}(\boldsymbol{y})\}$ для выбора параметров M, n **можно предложить метод 2 с использованием**:

– **для $\mathbb{C}$-подхода:** соотношений

$$H_{appr}^{(\mathbb{C})} M^{-\frac{2}{d}} = \frac{L}{3};\ \sup_{\boldsymbol{y}\in X} \left| \sum_{k=K+1}^{\infty} \left[\int_{\boldsymbol{X}} f_{\xi}(\boldsymbol{z}) \psi^{(k)}(\boldsymbol{z})\ d\boldsymbol{z} \right] \psi^{(k)}(\boldsymbol{y}) \right| = \frac{L}{3};\ \frac{H_{stoch}^{(\mathbb{C},pr)}(K)}{\sqrt{n}} = \frac{L}{3};$$

– **для $\mathbb{L}_2$-подхода:** соотношений

$$2\left[H_{appr}^{(\mathbb{L}_2)}\right]^2 M^{-\frac{4}{d}} = \frac{L^2}{3};\quad \frac{\mathrm{mes}\, X \times \max\limits_{i=1,\dots,M} \mathbf{D}\left[\sum_{k=1}^{K} \psi^{(k)}(\boldsymbol{\xi}) \psi^{(k)}(\boldsymbol{y}_i)\right]}{n} = \frac{L^2}{3};$$

$$2\, \mathrm{mes}\, X \sup_{\boldsymbol{y}\in X} \left(\sum_{k=K+1}^{\infty} \left[\int_{\boldsymbol{X}} f_{\xi}(\boldsymbol{z}) \psi^{(k)}(\boldsymbol{z})\ d\boldsymbol{z} \right] \psi^{(k)}(\boldsymbol{y}) \right)^2 = \frac{L^2}{3}$$

для некоторых положительных констант $H_{appr}^{(\mathbb{C})}, H_{stoch}^{(\mathbb{C},pr)}(K), H_{appr}^{(\mathbb{L}_2)}$.

З А К Л Ю Ч Е Н И Е

В докладе показано, как конструкции и теория функциональных алгоритмов метода Монте-Карло могут быть перенесены на компьютерные алгоритмы экономичного приближения (с заданной точностью) неизвестной плотности распределения случайной величины по заданной выборке.

Соображения теории условной оптимизации рассматриваемых компьютерных алгоритмов, связанные с выбором количества узлов аппроксимационной сетки и необходимого подмножества выборочных значений для достижения заданного уровня погрешности за минимальное время вычислений, показывают целесообразность использования на практике многомерного аналога полигона частот.

ВОПРОСЫ К СЛУШАТЕЛЯМ

1. Насколько актуальна постановка Задачи 1? Как и где искать приложения для рассматриваемых вычислительных схем?

2. Какие еще непараметрические оценки в узлах аппроксимационной сетки (кроме использования ядерных и проекционных оценок) целесообразно рассмотреть?

СПАСИБО ЗА ВНИМАНИЕ!

vav@osmf.sscc.ru