Исследование рандомизированных *квази* Монте-Карло алгоритмов

Федорова Надежда Сергеевна, 522-я группа

Санкт-Петербургский Государственный Университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель — д.ф.-м.н. **С.М. Ермаков**
Рецензент — д.ф.-м.н. **В.Б. Мелас**

Санкт-Петербург
2009г.

Направления исследования

Исследование применения рандомизированного метода *квази* Монте-Карло (РКМК) к ряду задач :

- Вычисление многомерных интегралов.
- Решение систем линейных алгебраических уравнений (СЛАУ).

Решение специальной задачи:

- Использование стохастической аппроксимации в методе Галеркина и распараллеливание метода.

Задача вычисления интеграла. Краткое описание методов Монте-Карло *квази* Монте-Карло (КМК)

Требуется приближенно вычислить интеграл $J[f] = \int f(x)\mu(dx)$.

- **Метод Монте-Карло**

  Оценка $J[f]$: $S_N[f] = \frac{1}{N}\sum\limits_{j=1}^{N} f(x_j)$, где $x_j \sim \mu$, независимы.

  Ошибка $\quad R_N[f] = |J[f] - S_N[f]|, \quad R_N[f] \sim O(N^{-\frac{1}{2}})$.

- **Метод *квази* Монте-Карло**
  Отличие от МК: x_j — квазислучайные.
  Пример (*последовательность точек Холтона*):
  q — простое число.
  $k = \sum\limits_{j=0}^{m} a_j(k) q^j$, $a_j(k)$ — коэффициенты разложения k в q-ичной системе.

  k-ое число Холтона по основанию q: $\psi_q(k) = \sum\limits_{j=0}^{m} \frac{a_j(k)}{q^{j+1}}$.

  d-мерная точка Холтона по осн. $q = (q_1, ..., q_d) : (\psi_{q_1}(k), ..., \psi_{q_d}(k))$,
  $q_1, ..., q_d$ — взаимно простые.
  Ошибка $\quad R_N[f] \sim O(\frac{\ln^d N}{N}) \quad$ (d — кратность интеграла).

Проблемы метода КМК

- **Проблема оценки погрешности:**
  Пусть $f(X)$ имеет огр. в смысле Харди-Краузе вариацию V_f, X_1,X_2,... — последовательность квазислуч. точек, тогда справедливо нер-во Коксмы-Хлавки:
  $$|R_N[f]| \leq \frac{V_f}{N} \cdot D^*(f),$$
  где $D^*(f) = D^*(f; X_1, ..., X_N) = \sup_{A \in \mathrm{A}} \left| N(\frac{\#\{X_i \in A\}}{N} - vol(A)) \right|$.
  Вычисление правой части неравенства задача более трудная, чем вычисление интеграла.
- **Проблема коррелированности точек:**
  Метод невозможно применять к сложным системам.

РКМК

Рандомизация — **случайное** отображение квазислучайных точек, также дающее лучший порядок убывания погрешности, чем обычный МК.
Примеры работ:
–«Методы Монте-Карло и Квази Монте-Карло для решения систем линейных алгебраических уравнений» (Рукавишникова А.И., 2008).
–«On Array-RQMC for Markov Chains: Mapping Alternatives and Convergence Rates» (P. L'Ecuyer, 2008). $(R_N[f] \sim O(N^{-\frac{3}{4}}))$

Задача вычисления интегралов. Оценка погрешности КМК с помощью РКМК

- **Выбранный способ рандомизации**
  d-мерн. т. Холтона: $Y_1, ..., Y_L,\ \ Y_i \in R^d$.
  $\Xi = (\alpha_1, ..., \alpha_d)^{\mathrm{T}}$ — р.р. на $[0, 1]^d$.
  Тогда $X_i = \{Y_i + \Xi\},\ \ i = 1, ..., L$.

- **Оценка погрешности КМК**
  Пусть точное значение интеграла: $\overline{J} = \int\limits_{[0,1]^d} f(x)dx$.

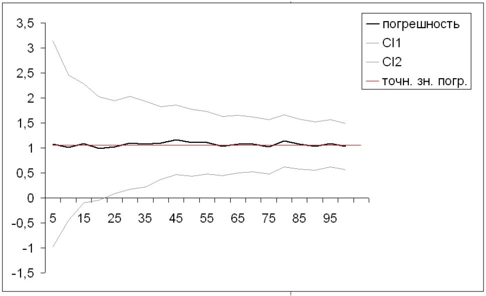
  $J_q = \frac{1}{L} \sum\limits_{m=1}^{L} f(Y_m)$ — оценка $\overline{J}$ методом КМК с т. Холтона по основанию q.
  Если имеется N независимых реализаций случайной величины
  $\xi = \frac{1}{L} \sum\limits_{k=1}^{L} f(X_k)$, то:

  $\widehat{J} = \frac{1}{N} \sum\limits_{p=1}^{N} \xi_p$ — вероятностная оц. $\overline{J}$, полученная с помощью метода РКМК.
  $R_N[f] = |\overline{J} - J_q|$ — погрешность метода КМК.
  Тогда $a = |\widehat{J} - J_q|$ — оценка для $R_N[f]$.

Задача вычисления интегралов. Оценка погрешности КМК с помощью РКМК. Пример

Пример

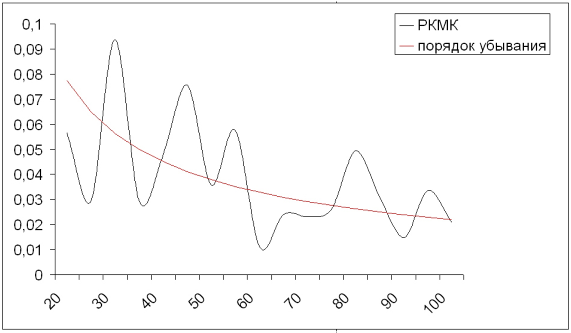
- *Тестовая функция* $f = \exp\{\sum i \cdot x_i\}/(1 + \prod \sin x_i/2)$.
- *Считаем* $\int_{[0,1]^d} f(x)dx$, $d = 3$, $q = \{13, 17, 19\}$.
- *Точное значение погрешности 1,048607.*



Зависимость оценки погрешности КМК от числа рандомизированных оценок.

Задача вычисления интегралов. Порядок убывания погрешности РКМК. Пример

- РКМК: $R_N[f] \sim O(N^{(-\alpha)})$, где $\alpha = 0,78$ — эмпирическая оценка (подтверждает порядок $O(N^{(-\frac{3}{4})})$).



Зависимость погрешности РКМК от числа рандомизированных оценок.

Задача решения СЛАУ

- Решается методом сеток $\Delta u = 0$ при заданных граничных условиях $u|_{\Gamma}$ в гиперкубе $[0,1]^d$.
- Разностные схемы: решение СЛАУ для нахождения $u(x_0)$, $x_0 \in [0,1]^d$.
- Статистический подход: ищем $u(x_0)$ с помощью случайных блужданий по равномерной сетке $S([0,1]^d)$ из x_0 до границы.

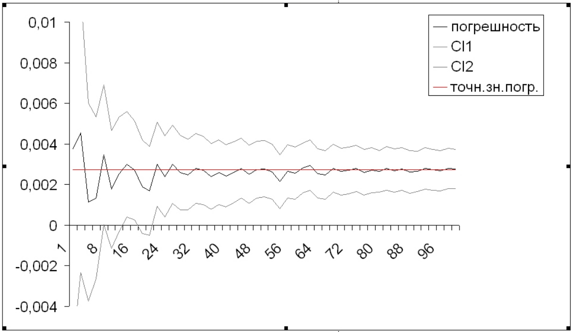
Оценка значения функции в точке: $u(x_0) \approx \frac{1}{N} \sum\limits_{x_i \in \Gamma} u(x_i)$.

- *Организация блуждания:*
  $x_1, x_2 ...$ — последовательность одномерных точек (МК, КМК, РКМК).
  d — размерность $\Rightarrow 2d$ — число направлений блуждания.
  Отрезок $[0,1]$ делится на $2d$ частей.
  i-ая точка попадает в одну из частей $[0,1]$, тем самым задавая направление блуждания на i-ой итерации.

Задача решения СЛАУ. Оценка погрешности КМК с помощью РКМК

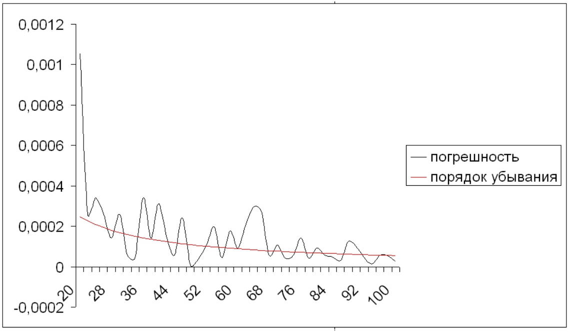
Пример

- *Тестовая функция:* $u = 1 + a \prod_{i=1}^{d} x_i.$
- *Размерность* $d = 10$, *параметр* $a = 4$, *область — единичный гиперкуб* $[0, 1]^d$.
- *Точное значение погрешности 0,00274201.*



- Исследованы оценки РКМК с разными основания т. Холтона, а также оценки МК. Во всех случаях $\sigma > \sigma_q$ (кроме случая $q = 5$).

- РКМК: $R_N[f] \sim O(N^{(-\alpha)})$, где $\alpha = 0,86$ — эмпирическая оценка (подтверждает порядок $O(N^{(-\frac{3}{4})})$).



Зависимость погрешности РКМК от числа рандомизированных оценок.

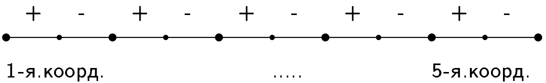
Задача решения СЛАУ.Средняя погрешность по области. Эффект «резонанса»

- Среднее значение погрешности по всему множеству точек сетки (размерность $d = 3$).

$$\Delta_{q;av} = \frac{1}{K} \sum_{x \in S([0,1]^d)} \Delta_q(x), \qquad K = \#\{x \in S([0,1]^d)\}$$

Δ_{av}	$\Delta_{3;av}$	$\Delta_{5;av}$	$\Delta_{7;av}$	$\Delta_{19;av}$	$\Delta_{29;av}$
0,1461072	0,369896	0,134548	0,09464	0,118794	0,1313664

- $\Delta_{av} > \Delta_{q;av}$ кроме случая $q = 3$.
- В рассмотренных примерах эффект $\sigma < \sigma_q$ и $\Delta_{av} < \Delta_{q;av}$ наблюдается в случаях, когда d и q не являются взаимнопростыми.
- Пример: $d = 5$, $q = 5$.



Метод Галеркина. Основные понятия

- Решается д.у. $L[y(x)] = f(x)$, заданы $y(x)|_{\Gamma}$ (L — дифференциальный оператор, может содержать частные или полные производные искомой функции).

- Решение ищется в виде $y(x) = y_0(x) + \sum\limits_{k=1}^{n} a_k y_k(x)$.

  $\{y_i(x)\}_{i=0}^{\infty}$ — полная ортогональная система функций.

  $y_0(x)|_{\Gamma} = y(x)|_{\Gamma}$

  $y_i(x)|_{\Gamma} = 0, \quad (0 < i < \infty)$

- Требование ортогональности $L[y(x)] - f(x)$ к функциям $y_i(x)$ дает систему

$$\int\limits_a^b [L(y(x)) - f(x)] y_k(x) dx = 0, \quad 1 \leq k \leq n.$$

- Получаем СЛАУ для коэффициентов a_k. Коэффициенты системы неявно выражены через интегралы $\Rightarrow$ с ростом n решение системы становится трудоемким.

Метод Галеркина. Подходы к распараллеливанию

- Область д.у. Ω разделяется на подобласти Ω_j. Ищется приближенное решение для Ω_j. «Обмен информацией» между граничащими областями . Минусы подхода: трудоемко и трудно реализуемо.
- Сначала подсчет интегралов, а потом решение полученной системы. Минусы подхода: трудоемко и плохо распараллеливается.
- Стохастическая аппроксимация («Random processes for classical equations of mathematical physics.» Ermakov S.M.,Nekrutkin V.V., Sipin A.S., 1989).

$AX = B$. $\{A_n, B_n\}$ — посл-ть нез. оценок $\{A, B\}$, все компоненты A_n и B_n имеют конечные дисперсии.
$X_0 \in R^d$ — константа, и пусть $2\alpha\lambda > 1$ (λ — наим. с.ч. A).
Процедура стохастической аппроксимации:

$$X_{n+1} = X_n + \frac{\alpha}{n}(B_{n+1} - A_{n+1}X_n),$$

Окончательная оценка: $\widehat{X} = \frac{1}{m}\sum\limits_{i=1}^{m}\widehat{X_i}$, где $\widehat{X_i}$ — независимые оценки решения, получаемые с помощью стохастической аппроксимации.

Теорема

В вышеописанных условиях, если элементы A_n и B_n имеют конечные третьи моменты, то распределение случайного вектора $\sqrt{n}(X_n - X^0)$ слабо сходится при $n \to \infty$ к нормальному распределению с нулевым средним значением.

Метод Галеркина. Стохастическая аппроксимация. Результаты

- Ур-е $y\prime\prime + y\prime = -2x$, $n = 3$, итерации $N = 1000$, число приближений $k = 15$, метод МК.
  Сравнение точных знач. коэфф. решения a_i с их оценками:

i	a_i	$\widehat{a_i}$	σ
1	0.4186	0.4264	0.0149
2	0.2035	0.1664	0.0195
3	-0.0407	-0.0508	0.0107

- Ур-е $y\prime\prime + y\prime = -2x$, $n = 10$, сравнение дисперсий методов МК и РКМК.
  - Оценка интеграла через одно случ. число: $\int f(X)dX \approx f(\alpha)$.
  - Оценка интеграла через несколько случ. чисел: $\int f(X)dX \approx \frac{1}{s}\sum_{i=1}^{s} f(\alpha_i)$.

i	$\widehat{a_i}$	$\widehat{a_{i5}}$	$\widehat{a_{i13}}$	$\widehat{a_{i29}}$
1	0,4306	0,4187	0,4250	0,3918
8	0,0014	0,0109	0,0372	0,0062

i	σ_i	σ_{i5}	σ_{i13}	σ_{i29}
1	0,0120	0,0180	0,0038	0,0145
8	0,0640	0,0764	0,0743	0,0101

Если сравнить дисперсии методов МК и РКМК, то в случае метода Галеркина однозначно нельзя сказать, что какой-то из методов "лучше".

Итоги работы

- В задачах вычисления интеграла и решения СЛАУ на примерах показана эффективность рандомизации для оценки погрешности метода КМК.
- Решение СЛАУ методами РКМК проводилось впервые, все результаты новые. Был обнаружен и объяснен эффект "резонанса".
- Показано, что погрешность метода РКМК ($O(N^{-\frac{3}{4}})$) лучше, чем в методе МК ($O(N^{-\frac{1}{2}})$), что также отмечалось в ряде работ.
- Указано, что для распараллеливания метода Галеркина может быть эффективным использование результатов, связанных со стохастической аппроксимацией.
- В методе Галеркина применение метода РКМК не дало ожидаемых результатов, что свидетельствует об определенных ограничениях в применении метода РКМК.