Две задачи теории метода Монте-Карло для решения уравнений баланса в пространстве мер

Румянцев Николай Алексеевич, 522-я группа

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

Научный руководитель — к.ф.-м.н. **В.В. Некруткин**
Рецензент — к.ф.-м.н. **Н.Э. Голяндина**

СтатМод

Санкт-Петербург
2008г.

Тематика, техника, литература. Задачи.

ТЕМАТИКА: Вероятностное решение уравнений баланса в пространстве мер.

ТЕХНИКА: (n, k)-частичные марковские скачкообразные процессы.

ЛИТЕРАТУРА: N. Golyandina and V. Nekrutkin, 1999.

- Анализ смещения и дисперсии оценок при больших n;
- Идея введения искусственного взаимодействия.

ДВЕ ЗАДАЧИ ДИПЛОМА:

- Уточнение вида смещения при $k > 2$;
- Изучение трудоемкости алгоритмов с искусственным взаимодействием в линейных задачах.

N. Golyandina and V. Nekrutkin

- УРАВНЕНИЕ БАЛАНСА:

$$d\mu_t/dt = B(\mu_t) - \mu_t, \quad \mu_t|_{t=0} = \mu \in H.$$

  H — распределения в $(D, \mathcal{B})$, $B(\nu) \in H$ при $\nu \in H$.
  Есть условия на D, B.

- ФУНКЦИОНАЛ: $\psi(\mu_t)$, $\psi \in \mathbf{C}^2(H)$.
- ПРОЦЕСС: марковский скачкообразный, n частиц движутся в D.
  - **Начальное положение:**
    n независимых с. в., имеющих распределение μ;
  - **Время между скачками:**
    показательное с $\lambda = n/k$;
  - **Скачок:**
    a) выбор k частиц из n;
    b) эти частицы совершают скачок согласно переходной функции B_k.

РЕЗУЛЬТАТ:
a) смещение и дисперсия имеют порядок $1/n$ (выбор B_k),
b) коэффициенты сложные (явный вид).

Уточнение вида смещения

БЫЛО:
Главный член смещения: kc_1/n.

ПРЕДЛОЖЕНИЕ:
На самом деле: $2c_1/n + (k-2)c_2/n$ при $k > 2$.

Замечание

- *Коэффициенты сложные (явный вид);*
- *При $k = 2$ — то же самое;*
- *При $k > 2$ есть примеры, когда $c_1 \equiv c_2$ и когда $c_1 \not\equiv c_2$;*
- *На среднеквадратическом отклонении поправка не сказывается.*

Линейная задача

ЛИНЕЙНАЯ ЗАДАЧА:

- **Линейное уравнение**

$$d\mu_t/dt = \int_D T(\,\cdot\,; u)\mu_t(du) - \mu_t = B(\mu_t) - \mu_t, \quad \mu_t|_{t=0} = \mu \in H,$$

  где $T(\,\cdot\,; u) \in H$ для всех $u \in D$;
- **Линейный функционал** $J(t) = \psi(\mu_t) = \int_D g d\mu_t,\ g \in \mathbf{C}(D)$.

ИЗВЕСТНО: Если

- $\zeta(t)$ — скачкообразный марковский процесс в D;
- начальное распределение μ;
- инфинитезимальные характеристики $\lambda = 1$ и $T(\,\cdot\,, u)$,

то

$$\mathsf{E} g(\zeta(t)) = J(t).$$

Обратное уравнение Колмогорова.

S-процесс

ОПРЕДЕЛЕНИЕ: n независимых копий процесса $\zeta(t)$. Обозначение — $\mathbf{S}$.

ОЦЕНКА: $\omega_n^{(\mathbf{s})} = \omega_n(g,t)$ — выборочное среднее (несмещенная).

Стандартная процедура.

ДИСПЕРСИЯ: $\mathrm{D}\big(\omega_n^{(\mathbf{s})}\big) = \sigma_S^2/n$.

ДРУГОЕ ОПИСАНИЕ: совместное движение n независимых копий процесса $\zeta(t)$.

ТРУДОЕМКОСТЬ моделирования на одну частицу:

$$L_{\mathbf{S}} = (t\Im_{\mathbf{S}} + \Im_0)\,\sigma_S^2.$$

$\Im_{\mathbf{S}}$ — затраты на один скачок процесса;
$\Im_0$ — затраты на начальное распределение.

Альтернатива: процесс $\mathbf{B}_\varphi^{(k)}$

ОПИСАНИЕ:

Случайная величина ω и функция $\varphi(\omega, u)$ такие, что $\mathcal{L}(\varphi(\omega, u)) = T(\,\cdot\,; u)$.

ПРОЦЕСС:

- **Частота скачков:** n/k;
- **Скачок:**
  - $(u_1, \ldots, u_n)$ — положение перед скачком;
  - выбирается k частиц;
  - их координаты $u_{i_1}, \ldots, u_{i_k}$ изменяются по правилу $u'_{i_j} \leftarrow \varphi(\omega, u_{i_j})$.

ВАЖНО: ω одно и то же для всех частиц!

Предложение

1. Несмещенная оценка.

2. Дисперсия: $n\mathrm{D}\big(\omega_n^{(\varphi)}\big) = \sigma_S^2 + (k-1)V_{\mathbf{B}_\varphi}(\psi, t, \mu) + o(1)$.

$\mathbb{V} = V_{\mathbf{B}_\varphi}/\sigma_{\mathbf{S}}^2$ — основная дисперсионная характеристика.

Процесс $\mathbf{B}_{\varphi}^{(k)}$: обсуждение

СРАВНЕНИЕ:

- дисперсии: $\mathsf{D}\big(\omega_n^{(\mathbf{s})}\big) \le \mathsf{D}\big(\omega_n^{(\varphi)}\big) \le k\mathsf{D}\big(\omega_n^{(\mathbf{s})}\big)$; $(0 \le \mathbb{V} \le 1)$.
- число скачков: у $\mathbf{B}_{\varphi}^{(k)}$ в k раз меньше;
- моделирование скачка: у $\mathbf{B}_{\varphi}^{(k)}$ — дополнительные операции.

Идея: если эти операции быстрые, то трудоемкость $\mathbf{B}_{\varphi}^{(k)}$ м. б. меньше, чем $\mathbf{S}$.

ОТНОШЕНИЕ ТРУДОЕМКОСТЕЙ:

$$L_{\mathbf{B}_{\varphi}}(k)/L_{\mathbf{S}} = \frac{t\Im_{\mathbf{B}}/k + \Im_0}{t\Im_{\mathbf{S}} + \Im_0}\Big(1 + (k-1)\mathbb{V}\Big),$$

- $\Im_{\mathbf{B}}$ и $\Im_{\mathbf{S}}$ — затраты на одно столкновение для процессов.
- $\Im_0$ — затраты на начальное распределение.

Процесс $\mathbf{B}_\varphi^{(k)}$: оптимальное k

ЗАТРАТЫ (вычисляются таймированием):

- $c(\omega) = c(T)$ — моделирование ω; $c(\lambda)$ — моделирование момента скачка;
- $c(\varphi)$ — вычисление φ; $kc(p)$ — выбор k координат из n.

ЗАТРАТЫ НА СКАЧОК на одну частицу:

$\mathbf{S}$-процесс: $\Im_\mathbf{S} = c(T) + c(\lambda)$;

$\mathbf{B}_\varphi^{(k)}$-процесс: $\Im_\mathbf{B} = c(T) + c(\lambda) + k(c(\varphi) + c(p))$.

ОПТИМАЛЬНОЕ k для $\mathbf{B}_\varphi^{(k)}$-процесса.

$$k_{opt}^2 \approx \frac{c(T) + c(\lambda)}{c(\varphi) + c(p) + \Im_0/t} \left(\frac{1}{\mathbb{V}} - 1\right).$$

МАКСИМАЛЬНОЕ отношение трудоемкостей: при $t \to \infty$ и $\mathbb{V} \to 0$

$$\frac{L_\mathbf{S}}{L_{\mathbf{B}_\varphi}(k_{opt})} \to \mathcal{C} = \frac{c(T) + c(\lambda)}{c(\varphi) + c(p)}.$$

Процесс $\mathrm{B}_\varphi^{(k)}$: примеры

УРАВНЕНИЕ: $D = \mathbb{R}$; $T(\,\cdot\,; u) = \mathbf{N}(u, 1)$, $\mu_0 = \mathbf{N}(0, 4)$; $\mu_t = \mathbf{N}(0, t + 4)$.

ВЗАИМОДЕЙСТВИЕ: $\omega \in \mathbf{N}(0, 1)$, $\varphi(\omega, u) = \omega + u$.

ЗАТРАТЫ (моделирование): $c(\varphi) = 1$, $c(T) = 18$, $c(p) = 10$ и $c(\lambda) = 16$.

ПРИМЕРЫ.

- Функционал: $\psi_1(\mu_t(\mu)) = \int_{\mathbb{R}} u^2\, \mu_t(du) = t + 4$;
- Функционал: $\psi_2(\mu_t(\mu)) = \int_{\mathbb{R}} \cos u\, \mu_t(du) = e^{-\left(1-1/\sqrt{e}\right)t-2}$.

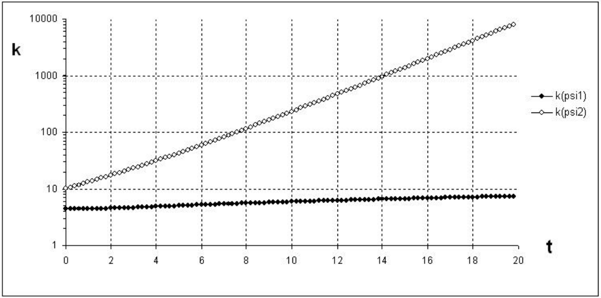

Рис.: Оптимальные k в зависимости от t (логарифмическая шкала).

Сравнение $\mathbf{B}_\varphi^{(k)}$ и S-алгоритмов: Теория

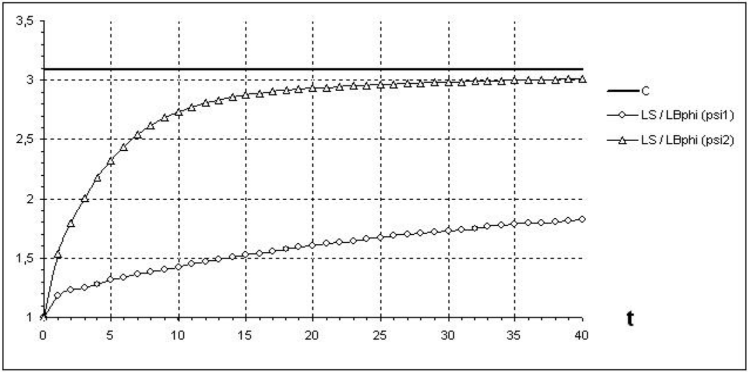

Рис.: Отношения $L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}(k_{opt})$ для функционалов ψ_1 и ψ_2.

Обе функции стремятся к $\mathcal{C} = 34/11 \approx 3.1$.

Сравнение $\mathbf{B}_{\varphi}^{(k)}$ и $\mathbf{S}$-алгоритмов: Моделирование

ПРИНЦИПЫ СРАВНЕНИЯ.

Для фиксированных k и t :

- Есть аналитические формулы для обеих дисперсий $\sigma^2_{B_\varphi}$ и σ^2_S (проверка моделированием).
- Подбор числа частиц n_{B_φ} и n_S — обратно пропорционально дисперсиям. Формула:

$$n_{B_\varphi} = \Big(1 + (k-1)\mathbb{V}\Big) n_S.$$

- N раз моделируются оба процесса.
- Сравниваются средние времена моделирования.

В экспериментах: n_S от 10^3 до 10^4 (в зависимости от k), N — от 10^4 до $2 \cdot 10^4$.

Сравнение $\mathbf{B}_{\varphi}^{(k)}$ и $\mathbf{S}$-алгоритмов: теория и таймирование.

ВОПРОС: насколько теоретическая трудоемкость соответствует таймированию?

Таблица: Функционал ψ_1. Отношение трудоемкостей и средних времен таймирования.

t	5	10	15
k_{opt}	5	6	7
$L_{\mathbf{S}}/L_{\mathbf{B}_{\varphi}}$	1.31	1.43	1.53
$\tau_S/\tau_{B_{\varphi}}$	1.41	1.45	1.51

Таблица: Функционал ψ_2. Отношение трудоемкостей и средних времен таймирования.

t	3	7
k_{opt}	23	82
$L_{\mathbf{S}}/L_{\mathbf{B}_{\varphi}}$	2.00	2.54
$\tau_S/\tau_{B_{\varphi}}$	2.23	2.70

ВЫВОД: хорошее соответствие.

Сравнение $B_\varphi^{(k)}$ и S-алгоритмов: оптимальное k.

ВОПРОС: соответствие оптимальных k

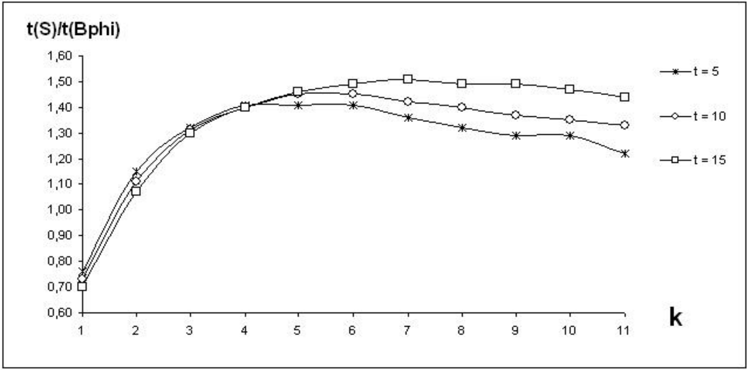

Рис.: Функционал ψ_1. Отношение τ_S/τ_{B_φ} при разных k.

Для ψ_2 аналогично (немного похуже).

Вторая альтернатива: процесс $\mathbf{S}_\varphi^{(k)}$. Сравнение с $\mathbf{S}$ и $\mathbf{B}_\varphi^{(k)}$

ОПРЕДЕЛЕНИЕ:

- m независимых реализаций (k,k)-частичного процесса;
- у каждой скачок как у $\mathbf{B}_\varphi^{(k)}$-процесса.

ОЦЕНКА : выборочное среднее (несмещенная).

Преимущество над $\mathbf{B}_\varphi^{(k)}$: Доверительные интервалы

ПРИНЦИПЫ СРАВНЕНИЯ (только таймирование).

- Выбор числа частиц аналогично предыдущему;
- Для $\mathbf{S}_\varphi^{(k)}$ замена дисперсии выборочной.

РЕЗУЛЬТАТЫ: функционал ψ_1.

- $k_{opt}(5) = 5$, трудоемкость в 1.3 раза меньше $\mathbf{S}$, совпадает с $\mathbf{B}_\varphi^{(k)}$;
- $k_{opt}(15) = 3$, трудоемкость в 1.1 раза меньше $\mathbf{S}$, в 1.5 раза больше $\mathbf{B}_\varphi^{(k)}$.

Сравнение алгоритмов $\mathbf{S}$, $\mathbf{B}_\varphi^{(k)}$ и $\mathbf{S}_\varphi^{(k)}$. Графики

РЕЗУЛЬТАТЫ: Функционал ψ_2.

- $k_{opt}(3) = 20$, трудоемкость в 3.8 раз меньше $\mathbf{S}$, в 1.7 раза меньше $\mathbf{B}_\varphi^{(k)}$;
- $k_{opt}(7) = 30$, трудоемкость в 5 раз меньше $\mathbf{S}$, в 1.5 раза меньше $\mathbf{B}_\varphi^{(k)}$.

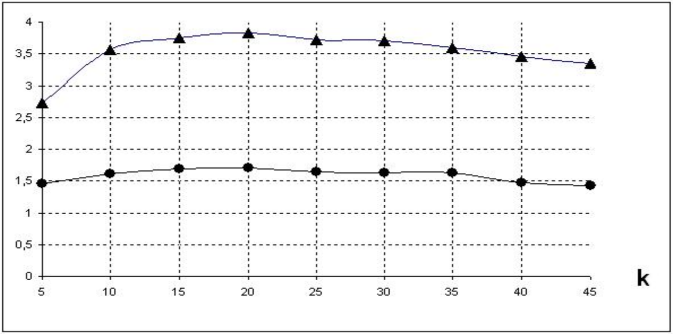

Рис.: Функционал ψ_2. Отношения τ_S/τ_{S_φ}(сверху) и $\tau_{B_\varphi}/\tau_{S_\varphi}$(снизу) при $t = 3$.

Выводы

ОБЩИЙ ВЫВОД: изучались 2 процесса с искусственным взаимодействием.

Перспективность процесса $\mathbf{B}_\varphi^{(k)}$ для больших времен, когда

- Моделирование процесса $\mathbf{S}$ стоит дорого, а взаимодействия — дешево. Формально,

$$\frac{c(T)+c(\lambda)}{c(\varphi)+c(p)} > 1.$$

- Дисперсия новой оценки не слишком большая. Формально,

$$\mathbb{V} = \frac{1}{k-1}\left(\frac{\sigma^2_{\mathbf{B}}}{\sigma^2_S} - 1\right) \quad \text{мало.}$$

Процесс $\mathbf{S}_\varphi^{(k)}$:

- Возможность построения доверительных интервалов.
- Недостаток теории.
- Качественно похоже на $\mathbf{B}_\varphi^{(k)}$ (эксперименты).