Спектральная метрика и спектральный тест для некоторых генераторов случайных чисел

Сабитов Рамис Рафаэлевич, 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**

Санкт-Петербург
2007г.

Постановка задачи: общая постановка

ОБЩАЯ ЗАДАЧА:

Изучение некоторых генераторов случайных чисел с помощью спектрального теста.

ТЕХНИКА:

Спектральная метрика.

Постановка задачи: теоретические основы

ИЗВЕСТНО:

- *Спектральная метрика*

$$\sigma(\mathcal{P}, \mathcal{Q}) = \sup_{\ell \in \mathbb{Z}^d,\, \ell \neq \mathbf{0}} | \varphi_{\mathcal{P}}(2\pi\ell) - \varphi_{\mathcal{Q}}(2\pi\ell) | / \|\ell\|_2,$$

  где $\mathcal{P}$ и $\mathcal{Q}$ — распределения, сосредоточенные на $\mathbb{I}^d = [0,1)^d$;

- σ — метрика на множестве распределений;

ИЗВЕСТНО:

Предложение

Пусть $\mathcal{P}_n, \mathcal{P}$ — распределения, сосредоточенные на единичном гиперкубе $\mathbb{I}^d = [0,1)^d$, причем $\mathcal{P}((0,1)^d) = 1$. Тогда для слабой сходимости $\mathcal{P}_n \Longrightarrow \mathcal{P}$ необходимо и достаточно, чтобы $\sigma(\mathcal{P}_n, \mathcal{P}) \to 0$.

Если U^d — равномерное распределение в $[0,1)^d$, то $w_d = 1/\sigma(\mathcal{P}, \mathrm{U}^d)$ — *волновое число распределения $\mathcal{P}$.*

Постановка задачи: изучение стохастических моделей генераторов

- Генераторы: МГ (мультипликативный); ЛКГ (линейный конгруэнтный) ; ЛКГПМ (линейный конгруэнтный с переменным мультипликатором).

  Всегда — периодическая последовательность $\{a_n\}_{n\geq 0}$ с некоторым периодом T;

- переход к стохастической модели генератора ($\{\delta_n\}_{n\geq 0}$):
  - рандомизация начального значения: δ_0 равномерна распределена на множестве H, где H — состав периода генератора $\{a_n\}_{n\geq 0}$;
  - рандомизация по периоду: $\delta_n = a_{n+\tau}$, где τ — равномерно распределена на множестве $\{0, \ldots, T-1\}$, T — период генератора;

  Обе рандомизации для МГ и ЛКГ — одно и тоже распределение, а для ЛКГПМ — нет.

  Для МГ, ЛКГ и ЛКГПМ при рандомизации по периоду $\{\delta_n\}_{n\geq 0}$ — стационарная последовательность.

  Для ЛКГПМ при рандомизации начального значения — нет.

Постановка задачи: изучение стохастических моделей генераторов

Если последовательность стационарная, то

$$\mathcal{P}^{(d)} \stackrel{\text{def}}{=} \mathcal{L}(\delta_0, \ldots, \delta_{d-1}) = \mathcal{L}(\delta_k, \ldots, \delta_{d-1+k})$$

для любого $k \in \mathbb{N}$ и задача состоит в изучении спектрального расстояния

$$\sigma(\mathcal{P}^{(d)}, \mathrm{U}_d) = \sup_{\ell \in \mathbb{Z}^d,\, \ell \neq \mathbf{0}} | \varphi^{(d)}(2\pi\ell) | / \|\ell\|_2,$$

где $\varphi^{(d)}$ — характеристическая функция распределения $\mathcal{P}^{(d)}$.

Нужно, чтобы $\sigma(\mathcal{P}^{(d)}, \mathrm{U}_d)$ было мало при всех d: $2 \leq d \leq d_0$, где d_0 задается априорно.

Мультипликативные генераторы

Мультипликативные генераторы: $a_{n+1} \stackrel{\text{def}}{=} \{\lambda a_n\}$, $a_0 = x_0/m$ исследовались при:

m — простое число, λ — первообразный элемент, $x_0 \neq 0$;

$m = 10^p$, x_0 — нечетное и не кратное 5;

$m = 2^p$.

ИЗВЕСТНО, что при $\lambda \equiv 5 \pmod 8$ и x_0 нечетном

$$\sigma(\mathcal{P}_p^{(d)}, \mathrm{U}_d) = \sup_{\ell \neq \mathbf{0}} \left\{ 1/\|\ell\|_2 : \ g(\ell) \equiv 0 \pmod{2^{p-2}} \right\},$$

где $g(\ell) = \sum_{s=0}^{d-1} l_s \lambda^s$. Есть эффективные алгоритмы вычисления.

Свойство волновых чисел:

$w_d^{(p)} \stackrel{\text{def}}{=} 1/\sigma(\mathcal{P}_p^{(d)}, \mathrm{U}_d) \leq \sqrt{\gamma_d}\ 2^{p-2}$ (неравенство Эрмита)

Нормированные волновые числа.

Мультипликативный генератор с $m = 2^p$ и $\lambda \equiv 3 \pmod 8$

Предложение

1. Если $\lambda \equiv 3 \pmod 8$ и x_0 — нечетное, то

$$\sigma(\mathcal{P}_p^{(d)}, \mathrm{U}_d) = \sup_{\ell \neq \mathbf{0}} \frac{1}{\|\ell\|_2} \begin{cases} 1/\sqrt{2}, & \text{если } g(\ell) \equiv 0 \pmod{2^{p-3}} \text{ и } g(\ell) \not\equiv 0 \pmod{2^{p-2}}, \\ 1, & \text{если } g(\ell) \equiv 0 \pmod{2^{p-1}}, \end{cases}$$

где $g(\ell) = \sum_{s=0}^{d-1} l_s \lambda^s$.

2. Пусть $p > 3$, $\lambda \equiv 5 \pmod 8$, $\lambda < 2^{p-2}$ и $\lambda^* = 2^{p-2} - \lambda$. Тогда

$$\sigma(\mathcal{P}_p^{(d)}(\lambda), \mathrm{U}_d) \geq \sigma(\mathcal{P}_{p+1}^{(d)}(\lambda^*), \mathrm{U}_d) \geq \sigma(\mathcal{P}_{p+2}^{(d)}(\lambda), \mathrm{U}_d).$$

3. Пусть $\lambda \equiv 3 \pmod 8$ и $\gamma_d^* = \gamma_d 2^{2/d}$. Тогда

$$w_p^{(d)}(\lambda) \leq \sqrt{\gamma_d^*}\, 2^{(p-2)/d}.$$

Есть алгоритм оценки волновых чисел сверху и снизу. Примерно в 80% случаев — точные значения.

Таблица: Оптимальные мультипликаторы при $d = 2$.

m	λ_5	σ_5	nw_5	λ_3	σ_3	nw_3
2^{16}	$5 + 8 \cdot 1261$	7.278e-3	9.988e-1	$3 + 8 \cdot 710$	6.470e-3	1.123e+0
2^{17}	$5 + 8 \cdot 1867$	5.197e-3	9.891e-1	$3 + 8 \cdot 958$	4.562e-3	1.126e+0
2^{18}	$5 + 8 \cdot 2900$	3.667e-3	9.910e-1	$3 + 8 \cdot 7229$	3.206e-3	1.133e+0
2^{19}	$5 + 8 \cdot 8521$	2.576e-3	9.978e-1	$3 + 8 \cdot 4004$	2.256e-3	1.139e+0

$$nw_5 = \frac{w_5}{\gamma_d^{1/2} m^{1/d}}, \quad nw_3 = \frac{w_3}{\gamma_d^{1/2} m^{1/d}};$$

Таблица: Оценки снизу для оптимальных волновых чисел, $d = 3$ и 4.

m	$w_5^{(3)}$	$w_3^{(3)} \geq$	$w_5^{(4)}$	$w_3^{(4)} \geq$
2^{14}	1.62e+01	**1.84e+01**	8.60e+00	**9.27e+00**
2^{15}	2.17e+01	**2.25e+01**	1.05e+01	**1.11e+01**
2^{16}	2.70e+01	**3.07e+01**	1.21e+01	**1.34e+01**
2^{17}	3.49e+01	**3.82e+01**	1.51e+01	**1.59e+01**

МГ: численные эксперименты - 2

Таблица: Нормированные волновые числа (L'Ecuyer). $2 \leq d \leq 8$, $m = 2^{32}$, $\lambda \equiv 5 (\mathrm{mod}\, 8)$

λ	$nw^{(2)}$	$nw^{(3)}$	$nw^{(4)}$	$nw^{(5)}$	$nw^{(6)}$	$nw^{(7)}$	$nw^{(8)}$	min
438293613	0.76	0.83	0.83	0.79	0.75	0.83	0.76	0.75
741103597	0.87	0.80	0.80	0.80	0.76	0.76	0.77	0.76

Полный перебор.

Таблица: Нормированные волновые числа. $2 \leq d \leq 8$, $m = 2^{32}$, $\lambda \equiv 3 (\mathrm{mod}\, 8)$

λ	$nw^{(2)}$	$nw^{(3)}$	$nw^{(4)}$	$nw^{(5)}$	$nw^{(6)}$	$nw^{(7)}$	$nw^{(8)}$	min
26788971	0.91	0.84	0.84	0.81	0.81	0.79	0.77	0.77
573964635	0.94	0.88	0.81*	0.88	0.81	0.83	0.77	0.77

Случайный перебор.

ВЫВОД: согласно спектральному тесту, случай $\lambda \equiv 3 \,(\mathrm{mod}\, 8)$ предпочтительнее.

Линейный конгруэнтный генератор

Линейный конгруэнтный генератор (ЛКГ):

$$a_{n+1} \stackrel{\text{def}}{=} \{\lambda a_n + c\},$$

где $a_0 = x_0/m$, $c = r/m$;

Для $m = 2^p$ максимальный период $T = 2^p$ при $\lambda \equiv 1 \pmod 4$, $r \neq 0$.

ИЗВЕСТНО, что при $\lambda \equiv 1 \pmod 4$

$$\sigma(\mathcal{P}_p^{(d)}, \mathrm{U}_d) = \sup_{\ell \neq \mathbf{0}} \{1/\|\ell\|_2 : \ g(\ell) \equiv 0 \pmod{2^p}\},$$

где $g(\ell) = \sum_{s=0}^{d-1} l_s \lambda^s$. Есть эффективные алгоритмы вычисления.

Линейный конгруэнтный генератор с переменным мультипликатором

Линейный конгруэнтный генератор с переменным мультипликатором (ЛКГПМ):

$$a_{n+1} \stackrel{\text{def}}{=} \begin{cases} \{\lambda_1 a_n + c_1\}, & \text{если} \quad n\text{—четное}, \\ \{\lambda_2 a_n + c_2\}, & \text{если} \quad n\text{—нечетное}, \end{cases}$$

где $a_0 = x_0/m$, $c_1 = r_1/m$, $c_2 = r_2/m$.

Предложение

Пусть $m = 2^p$, $\lambda_1 \equiv 1\,(mod\,4)$, $\lambda_2 \equiv 1\,(mod\,4)$.
Если r_1 *и* r_2 *имеют разную четность, то генератор обладает максимальным периодом* $T = 2^{p+1}$.
Состав периода — $\{0, 1 \ldots, 2^p - 1\}$.

ЛКГПМ: рандомизация начального значения

Последовательность $\{\delta_n\}_{n\geq 0}$ не стационарная: $(\delta_0, \ldots, \delta_{d-1})$, $(\delta_1, \ldots, \delta_d)$;

Предложение

1.

$$\sigma(\mathcal{P}_1^{(d)}, \mathrm{U}_d) = \max_{\ell \in \mathbb{Z}^d} \{1/\|\ell\|_2 : \ell \neq \mathbf{0},\, g_1(\ell) \equiv 0 \,(mod\, 2^p)\},$$

где $\mu = \lambda_1\lambda_2$ *и*

$$g_1(\ell) = \sum_{j=0}^{\lceil d/2-1 \rceil} \mu^j \ell_{2j} + \sum_{j=0}^{\lfloor d/2-1 \rfloor} \mu^j \lambda_1 \ell_{2j+1},$$

2. $\sigma(\mathcal{P}_2^{(d)}, \mathrm{U}_d)(\lambda_1, \lambda_2) = \sigma(\mathcal{P}_1^{(d)}, \mathrm{U}_d)(\lambda_2, \lambda_1)$.

Мера отклонения от равномерности: $\sigma_d = \max\{\sigma(\mathcal{P}_1^{(d)}, \mathrm{U}_d), \sigma(\mathcal{P}_2^{(d)}, \mathrm{U}_d)\}$.
Волновое число: $w_d = 1/\sigma_d$.
Алгоритм — такой же как у ЛКГ.

ЛКГПМ, рандомизация начального значения: численные эксперименты

Таблица: Примеры мультипликаторов для моделей ЛКГ и ЛКГПМ, $2 \le d \le 7$

№	λ	m	$nw^{(2)}$	$nw^{(3)}$	$nw^{(4)}$	$nw^{(5)}$	$nw^{(6)}$	$nw^{(7)}$	min
1	8749, 13413	2^{15}	0.84	0.75	0.76	0.77	0.75	0.75	0.75
2	1765, 19865	2^{15}	0.77	0.81	0.71	0.73	0.75	0.71	0.71
3	21049	2^{15}	0.92	0.91	0.75	0.72	0.73	0.75	0.72
4	8749	2^{15}	0.84	0.80	0.71	0.70	0.73	0.71	0.70
5	149, 16113	2^{14}	0.80	0.73	0.79	0.74	0.78	—	0.73
6	1085, 9425	2^{14}	0.80	0.74	0.77	0.74	0.75	—	0.74
7	2649	2^{14}	0.83	0.72	0.83	0.77	0.78	—	0.72
8	1325	2^{14}	0.96	0.71	0.78	0.72	0.75	—	0.71

ВЫВОД: ЛКГПМ с рандомизацией по начальному значению не хуже ЛКГ, но перебор очень большой.

ЛКГПМ: рандомизация по периоду

Предложение

1. Последовательность $\{\beta_n\}_{n\geq 0}$ *стационарная.*
2. Пусть $\lambda_1 = \lambda_2 = \lambda$, *тогда*

$$\sigma(\mathcal{P}^{(d)}, \mathrm{U}_d) = \max_{\ell\in\mathbb{Z}^d} \frac{1}{\|\ell\|_2} \left\{ \left| \cos\left(\pi a \, \frac{g(\ell) - A}{2^p(\lambda+1)} \right) \right| : \ell \neq \mathbf{0},\, g(\ell) \equiv 0 \,(mod\, 2^p) \right\},$$

где $g(\ell) = \sum_{k=0}^{d-1} \lambda^k l_k$, $a = r_2 - r_1$ *и*

$$A = \sum_{j=0}^{\lceil d/2-1 \rceil} \ell_{2j} - \sum_{j=0}^{\lfloor d/2-1 \rfloor} \ell_{2j+1}.$$

Лучше, чем при рандомизации по начальному значению для модели ЛКГ.