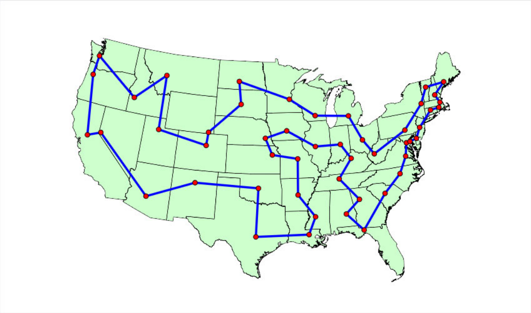
Анализ жадного алгоритма для задачи коммивояжера на случайных входных данных

Трегубов Даниил Валерьевич МЕН-470206

18 июня 2021 г.

В настоящей работе применяют следующие обозначения и сокращения.
ИБГ - Алгоритим Иди в Ближайщий Город
TSP - Travelling Salesman Problem, Задача коммивояжера
Greedy - вес маршрута коммивояжера, найденного в графе алгоритмом ИБГ
Optimal - вес оптимального маршрута коммивояжера в графе

Введение



Приложения задачи коммивояжера

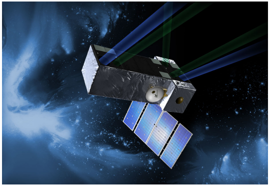
1. Посещение n городов, населенных пунктов, точек

Приложения задачи коммивояжера

1. Посещение n городов, населенных пунктов, точек
2. Устройство, которое может выполнять только одну работу

Приложения задачи коммивояжера

1. Посещение n городов, населенных пунктов, точек
2. Устройство, которое может выполнять только одну работу
3. Программа для наведения интерферометра



Математическая модель для задачи коммивояжера

Дано:

полный взвешенный граф $G = (V, E, w)$ без петель.

Найти:

гамильтонов цикл минимального веса

Постановки задачи коммивояжера

1. Классическая постановка

Постановки задачи коммивояжера

1. Классическая постановка
2. Асимметричная постановка

Постановки задачи коммивояжера

1. Классическая постановка
2. Асимметричная постановка
3. Метрическая постановка

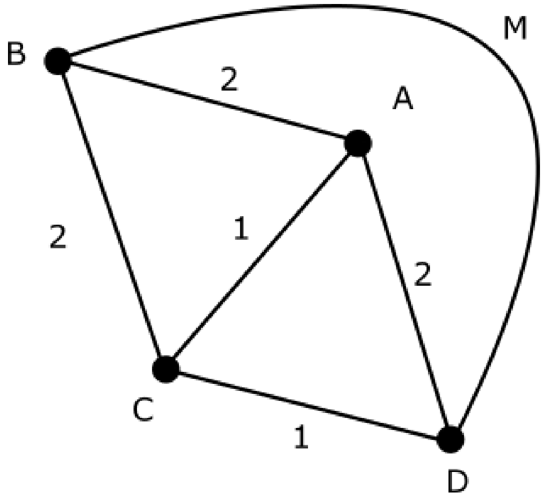
Постановки задачи коммивояжера

1. Классическая постановка
2. Асимметричная постановка
3. Метрическая постановка
4. Max TSP

Алгоритм ИБГ

Одним из алгоритмов, находящих некое решение за приемлемое количество операций, а именно $O(n^2)$, где n - количество вершин, является жадный алгоритм с выбором ближайшей вершины. Суть алгоритма заключается в том, что на каждом шаге мы просматриваем все непосещенные вершины в которые можно попасть из текущей вершины, и выбираем ту, расстояние до которой минимально.

Пример неоптимальной работы алгоритма ИБГ



Асимптотическая оптимальность

Алгоритм $\mathcal{A}$ будем называть асимптотически оптимальным если существуют такие $\epsilon_n \to 0, \delta_n \to 0$ при $n \to \infty$ что применение алгоритма $\mathcal{A}$ дает значение $\mathcal{Z}_{\mathcal{A}}$, удовлетворяющее неравенству

опр

$$P\{\mathcal{Z}_{\mathcal{A}} \leq (1+\epsilon_n)\mathcal{Z}\} \geq 1-\delta_n \tag{1}$$

где $\mathcal{Z}_{\mathcal{A}}$ - решение, найденное алгоритмом, $\mathcal{Z}$ - минмальное решение. Это определение при $\epsilon_n \equiv 0$ совпадает с понятием алгоритма, который почти всегда приводит к точному решению. Альтернативный вариант формулы из определения:

Вероятность несрабатывания

$$P\{\mathcal{Z}_{\mathcal{A}} > (1+\epsilon_n)\mathcal{Z}\} \leq \delta_n \tag{2}$$

Теорема Петрова

Пусть $X_1, \ldots, X_n$ — независимые случайные величины и существуют положительные постоянные $g_1, \ldots, g_n$ и T такие, что для всех $0 \leq t \leq T$

$$\mathbb{E}e^{tX_k} \leq e^{\frac{1}{2}g_k t^2} \tag{3}$$

Положим $S = \sum_{k=1}^{n} X_k$ и $G = \sum_{k=1}^{n} g_k$. Тогда

Теорема Петрова

$$P\{S > x\} \leq \begin{cases} exp(-\frac{x^2}{2G}), & 0 \leq x \leq GT \\ exp(-\frac{Tx}{2}), & x \geq GT \end{cases} \tag{4}$$

Общее условие асимптотической оптимальности

Найдем условия, при выполнении которых жадный алгоритм является асимптотически оптимальным, то есть

$$P\{\mathcal{Z}_{\mathcal{A}'} > (1+\epsilon_n)\mathcal{Z}\} \leq \delta_n \tag{5}$$

где $\epsilon_n \longrightarrow 0, \delta_n \longrightarrow 0 \; n \longrightarrow \infty$

Вес маршрута коммивояжера, полученного применением жадного алгоритма обозначим как $\mathcal{Z}_{\mathcal{A}'}$. Он определяется формулой

$$\mathcal{Z}_{\mathcal{A}'} = \sum_{k=1}^{n} a_{i_k i_{k+1}} \tag{6}$$

Вес маршрута отпимального решения задачи обозначим как $\mathcal{Z}$. Он определяется формулой:

$$\mathcal{Z} = \min_{\{\pi\}} \sum_{k=1}^{n} a_{i_k i_{k+1}} \tag{7}$$

Алгоритм $\mathcal{A}'$ является асимптотически оптимальным при выполнении условий $\lim\limits_{n\to\infty} \mathcal{I}_n = \infty$ и

$$\frac{b}{a} \leq \frac{1}{\Psi(n)} min\{\frac{1}{\gamma_n}, \frac{n}{\mathcal{I}_n}\} \tag{8}$$

где:

γ_n - корень уравнения $F(\gamma) = \frac{1}{n}$, то есть $\gamma_n = F^{-1}(\frac{1}{n})$

$\mathcal{I}_n = \int_{\gamma_n}^{1} \frac{dx}{F(x)}$,

$\Psi(n)$ - произвольная растущая от n функция, $\lim\limits_{x\to\infty} \Psi(n) = \infty$

Вероятностный анализ ИБГ для равномерного распределения

Определим условия асимптотической оптимальности алгоритма $\mathcal{A}$ для случая, когда веса ребер графа a_{ij} могут быть выбраны равновероятно из отрезка [a,b], a>0.

Теорема 2

Если элементы a_{ij} матрицы A принимают значения равновероятно из отрезка [a,b], то алгоритм $\mathcal{A}$ является асимптотически оптимальным при выполнении следующего условия:

$$\frac{b}{a} \leq \frac{n}{ln(n)} \frac{1}{\Psi(n)} \tag{9}$$

Пусть $\epsilon_n = \frac{c}{\Psi(n)}$

Возьмем $\Psi(n) = \sqrt{\frac{cn}{ln(n)}}$

$$\epsilon_n \leq \sqrt{\frac{cln(n)}{n}} \tag{10}$$

Вероятностный анализ ИБГ для показательного распределения

Теорема
Задача Random Min TSP в случае показательного распределения весов ребер на $[a_n, \infty)$ решается алгоритмом ИБГ за $O(n^2)$ с оценками относительной погрешности

$$\epsilon_n = 5\frac{\alpha_n/a_n}{n/ln(n)} \tag{11}$$

и вероятности несрабатывания

$$\delta_n = O(n^{-1}) \tag{12}$$

асимптотически точно при условии

$$\frac{\alpha_n}{a_n} = o(\frac{n}{ln(n)}) \tag{13}$$

доказательство

Оценим вероятность несрабатывания алгоритма

$$\begin{aligned} P\{W(\pi) > (1+\epsilon_n)OPT\} &\leq \\ \leq P\{S > \frac{n a_n \epsilon_n}{\alpha_n}\} &= \\ = P\{\tilde{S} > \frac{n a_n \epsilon_n}{\alpha_n} - ES\} & \end{aligned} \tag{14}$$

Оценим ES

$$\begin{aligned} ES = E(\sum_{k=1}^{n-1} c_{\pi_k,\pi_{k+1}} + c_{\pi_n,\pi_1}) = \\ = \sum_{k=1}^{n-1} E(X_1(k)) + E(c_{\pi_n,\pi_1}) \end{aligned} \tag{15}$$

где $X_1(k)$ - минимум из k случайных независимых переменных.

$$F_{X_1(k)} = 1 - (1 - F(x))^k = 1 - e^{-kx}, \ x \geq 0 \tag{16}$$

$$EX_1(k) = \int_0^\infty xd(F_{X_1(k)}) = \frac{1}{k} \tag{17}$$

Величина c_{π_n,π_1} распределена по показательному закону. Найдем матожидание:

$$Ec_{\pi_n,\pi_1} = Ep(x) = \int_0^{\infty} xe^{-x}dx = 1 \tag{18}$$

$$\begin{aligned} \sum_{k=1}^{n-1} E(X_1(k)) + E(c_{\pi_n,\pi_1}) \leq \\ \leq \sum_{k=1}^{n-1} \frac{1}{k} + 1 \leq \\ \leq ln(n) + C + 1 \end{aligned} \tag{19}$$

Где $C < 0.58$

$$ES \leq ln(n) + C + 1 \tag{20}$$

$$P\{\tilde{S} > \frac{na_n\epsilon_n}{\alpha_n} - ES\} \leq$$
$$P\{\tilde{S} > \frac{na_n\epsilon_n}{\alpha_n} - ln(n) - C - 1\} \qquad (21)$$

Подставим ϵ_n из условия.

$$\epsilon_n = 5\frac{\alpha_n/a_n}{n/ln(n)} \qquad (22)$$

По определению вероятности несрабатывания $\epsilon_n \to 0$ при $n \to \infty$ Данное условие выполняется, так как по условию теоремы $\frac{\alpha_n}{a_n} = o(\frac{n}{ln(n)})$.

$$P\{\tilde{S} > 5ln(n) - ln(n) - C - 1\} =$$
$$P\{\tilde{S} > 4ln(n) - C - 1\} \qquad (23)$$

Применим теорему Петрова при $x = 4ln(n) - C - 1$.
Нас интересует асимптотика, а значит большие значения H, поэтому мы можем рассмотреть только случай $x \geq HT$.

$$P\{\tilde{S} > x\} \leq e^{-\frac{Tx}{2}} \tag{24}$$

Лемма[1]. Пусть $T = \frac{1}{2\alpha_n}$ и $g_k = \frac{3\alpha_n^2}{k^2}$, $1 \leq k \leq n-2$. Тогда при любых k и t где $k = 1, .., n-2$ и $0 \leq t \leq T$, справедливо неравенство:

$$Ee^{t[\xi_k - E\xi_k]} \leq exp(\frac{g_k t^2}{2}) \tag{25}$$

Доказательство данной леммы означает, что для $T = \frac{1}{2\alpha_n}$, $X_i = \xi_k - E\xi_k$ и $h_i = \frac{3\alpha_n^2}{k^2}$ выполняется условие теоремы Петрова.
Подставим найденные значения в оценку асимптотики.
В нормированном случае $T = \frac{1}{2}$.

[1] Гимади Э.Х., Хачай М.Ю. Экстремальные задачи на множествах перестановок, Екатеринбург, 2016.

$$P\{\tilde{S} > x\} \leq e^{-\frac{x}{4}} \tag{26}$$

Подставим x.

$$P\{\tilde{S} > 4ln(n) - C - 1\} \leq e^{-\frac{4ln(n)-C-1}{4}} \leq e^{-ln(n)} \leq n^{-1} = O(n^{-1}) \tag{27}$$

Таким образом, мы доказали, что

$$P\{W(\pi) > (1+\epsilon_n)OPT\} = O(n^{-1}) \tag{28}$$

Вычислительный эксперимент

Тип распределения	Размерность графа	$1+\epsilon$	Превышений $1+\epsilon$
равномерное	5	2.28755032994	0
равномерное	10	1.9210340371976	0
равномерное	20	1.599146454710	0
равномерное	40	1.36888794541	0
равномерное	80	1.2191013317	0
равномерное	160	1.12687934538	0
равномерное	320	1.07210401244	0
показательное	5	2.6094379124341	2
показательное	10	2.151292546497	11
показательное	20	1.7489330683884	5
показательное	40	1.4611099317	2
показательное	80	1.27387666466	0
показательное	160	1.1585991817	0
показательное	320	1.09013001555	0

ЗАКЛЮЧЕНИЕ

Интерес для дальнейших исследований представляет поиск условий асимптотической оптимальности для других распределений. Также определенный интерес вызывает более точная оценка ϵ для случаев равномерного и показательного распределений.

Спасибо за внимание!

Найдем условия, при выполнении которых жадный алгоритм является асимптотически оптимальным, то есть

$$P\{\mathcal{Z}_{\mathcal{A}'} > (1+\epsilon_n)\mathcal{Z}\} \leq \delta_n \tag{29}$$

где $\epsilon_n \to 0, \delta_n \to 0$ $n \to \infty$
Вес маршрута коммивояжера, полученного применением жадного алгоритма обозначим как $\mathcal{Z}_{\mathcal{A}'}$. Он определяется формулой

$$\mathcal{Z}_{\mathcal{A}'} = \sum_{k=1}^{n} a_{i_k i_{k+1}} \tag{30}$$

Вес маршрута отпимального решения задачи обозначим как $\mathcal{Z}$. Он определяется формулой:

$$\mathcal{Z} = \min_{\{\pi\}} \sum_{k=1}^{n} a_{i_k i_{k+1}} \tag{31}$$

Оценим сверху левую часть (29). Так как $a > 0$, то $\mathcal{Z} \geq na$ и

$$P\{\mathcal{Z}_{\mathcal{A}} > (1+\epsilon_n)\mathcal{Z}\} \leq P\{\mathcal{Z}_{\mathcal{A}} > (1+\epsilon_n)na\} \tag{32}$$

Обозначим через$\mathcal{Z}^*_{\mathcal{A}'}$и $\mathcal{D}^*_{\mathcal{A}'}$верхние оценки соответственно математического ожидания $E(\mathcal{Z}_{\mathcal{A}'})$и дисперсии $D(\mathcal{Z}_{\mathcal{A}'})$случайной величины $\mathcal{Z}_{\mathcal{A}'}$. Обозначим $\Delta_n = (1+\epsilon_n)na - \mathcal{Z}^*_{\mathcal{A}'}$. Получаем:

$$P\{\mathcal{Z}_{\mathcal{A}'} > (1+\epsilon_n)na\} = P\{\mathcal{Z}_{\mathcal{A}'} > \mathcal{Z}^*_{\mathcal{A}'} + [(1+\epsilon_n)na - \mathcal{Z}^*_{\mathcal{A}'}]\} \leq P\{\mathcal{Z}_{\mathcal{A}'} > E(\mathcal{Z} \tag{33}$$

Пусть $\epsilon_n = K(\frac{\mathcal{Z}^*_{\mathcal{A}'}}{na} - 1)$, $K > 1$. Тогда $\Delta_n = (K-1)(\mathcal{Z}^*_{\mathcal{A}'} - na) \geq 0$. Продолжим неравенство (33) применив неравенство Чебышева.

$$P\{\mathcal{Z}_{\mathcal{A}'} > E(\mathcal{Z}_{\mathcal{A}'}) + \Delta_n\} \leq P\{|\mathcal{Z}_{\mathcal{A}'} - E(\mathcal{Z}_{\mathcal{A}'}))| \geq \Delta_n\} \leq \frac{D(\mathcal{Z}_{\mathcal{A}'})}{\Delta_n^2} \leq \frac{\mathcal{D}^*_{\mathcal{A}'}}{\Delta_n^2} = \frac{\mathcal{D}^*_{\mathcal{A}'}}{(K-1)^2(\mathcal{Z}^*_{\mathcal{A}'} - na)^2} \tag{34}$$

Так как K - константа, $K > 1$, то из данной цепочки неравенств следует, что условие ассимптотической оптимальности жадного алгоритма будет выполнено если мы покажем, что $\epsilon_n = K(\frac{\mathcal{Z}^*_{\mathcal{A}'}}{na} - 1) \to 0$ и $\delta_n = \frac{\mathcal{D}^*_{\mathcal{A}'}}{(K-1)^2(\mathcal{Z}^*_{\mathcal{A}'} - na)^2} \to 0$ при $n \to \infty$

Вычисление верхних оценок $\mathcal{Z}^*_{\mathcal{A}'}$и $\mathcal{D}^*_{\mathcal{A}'}$

Математическое ожидание $E(\mathcal{Z}_{\mathcal{A}'})$равно сумме матожиданий величин $a_{i_k i_{k+1}}$ минимальных длин ребер, ведущих в еще непосещенную вершину на k-м шаге алгоритма $\mathcal{A}'$. В целях удобства дальнейших вычислений пронормируем случайную величину ξ значений элементов a_{ij} длин ребер графа, положив $\xi' = \frac{\xi - a}{b - a}$. Обозначим через l_kзначение матожидания нормированной случайной величины $l_k = a'_{i_k i_{k+1}}$. На k-ом шаге алгоритма выбирается минимум из $n - k$ элементов. В силу независимости этих элементов вероятность $\Phi_k(x)$того, что величина l_kминимального из этих элементов не превышает величины x равна

$$\Phi_k(x) = P\{l_k \leq x\} = 1 - (1 - F(x))^{n-k} \tag{35}$$

, где

$$F_k(x) = P\{\xi' \leq x\}, 0 \leq x \leq 1 \tag{36}$$

Тогда величина $E(l_k)$ равна

$$E(l_k) = \begin{cases} \int_0^1 x d\Phi_k(x), k = 1, 2, ...n-1 \\ E(l_{n-1}), k = n \end{cases} \qquad (37)$$

откуда получим

$$E(l_k) = x\Phi_k(x)|_0^1 - \int_0^1 \Phi_k(x)dx =$$
$$= 1 - \int_0^1 [1 - (1 - F(x))^{n-k}]dx = \int_0^1 (1 - F(x))^{n-k}]dx \qquad (38)$$
$$k = \overline{1, n-1}$$

В силу нормировки минимальный элемент $a_{i_{k-1} i_k}$ связан с величиной l_kсоотношением $a_{i_k i_{k+1}} = a + (b-a)l_k$. Поэтому

$$E(\mathcal{Z}_{\mathcal{A}'}) = \sum_{k=1}^{n} [a + (b-a)E(l_k)] \qquad (39)$$

.

Откуда с учетом (38) имеем:

$$E(\mathcal{Z}_{\mathcal{A}'}) = na + (b-a)[\int_0^1 \sum_{k=1}^{n-1} (1-F(x))^{n-k} dx + \int_0^1 (1-F(x))dx] =$$

$$= na + (b-a)\int_0^1 \frac{[1-F(x)][1-(1-F(x))^{n-1}+F(x)]}{F(x)} dx \le$$

$$\le na + (b-a)\int_0^1 \frac{1-(1-F(x))^n}{F(x)} dx =$$

$$= na + (b-a)[\int_0^{\gamma_n} \frac{1-(1-F(x))^n}{F(x)} dx + \int_{\gamma_n}^1 \frac{1-(1-F(x))^n}{F(x)} dx] \tag{40}$$

где γ_n - корень уравнения $F(\gamma) = \frac{1}{n}$, то есть $\gamma_n = F^{-1}(\frac{1}{n})$.
Учитывая что при $0 \le Z \le 1$ справедливо неравенство $\frac{1-(1-Z)^n}{Z} \le n$, оценку $E(\mathcal{Z}_{\mathcal{A}})$ можем продолжить следующим образом

$$E(\mathcal{Z}_{\mathcal{A}'}) \le \mathcal{Z}^*_{\mathcal{A}'} = na + (b-a)[\gamma_n * n + \int_{\gamma_n}^1 \frac{dx}{F(x)}] \tag{41}$$

Перейдем к вычислению верхней оценки $\mathcal{D}^*_{\mathcal{A}}$. Дисперсия d_k случайной нормированной величины l_k на k-ом шаге равна

$$d_k = \int_0^1 (x - E(l_k))^2 d\Phi_k(x) = \int_0^1 x^2 d\Phi_k(x) - (E(l_k))^2 < \int_0^1 x d\Phi_k(x) = \tag{42}$$

Тогда с учетом того, что дисперсия минимального элемента $a_{i_k i_{k+1}}$ равна $(b-a)^2 d_k$, дисперсия случайной величины $\mathcal{Z}_{\mathcal{A}}$ с учетом (39) оценивается следующим образом:

$$\begin{gathered} \mathcal{D}(\mathcal{Z}_{\mathcal{A}}) = \sum_{k=1}^{n} (b-a)^2 d_k < (b-a)^2 \sum_{k=1}^{n} E(l_k) = \\ = (b-a)^2 * \frac{E(\mathcal{Z}_{\mathcal{A}}) - na}{(b-a)} \leq (b-a)(\mathcal{Z}^*_{\mathcal{A}} - na) \end{gathered} \tag{43}$$

Окончательно с учетом (41) получаем верхнюю оценку для $\mathcal{D}(\mathcal{Z}_{\mathcal{A}})$

$$\mathcal{D}(\mathcal{Z}_{\mathcal{A}}) < \mathcal{D}^*_{\mathcal{A}} = (b-a)^2 [\gamma_n * n + \int_{\gamma_n}^1 \frac{dx}{F(x)}] \tag{44}$$

Вернемся к неравенству (34). Имея в виду полученные оценки (41) и (44), выражения для ϵ_n и δ_n можно записать в следующем виде:

$$\epsilon_n = K(\frac{b}{a} - 1)[\gamma_n + \frac{1}{n}\int_{\gamma_n}^{1} \frac{dx}{F(x)}] \tag{45}$$

$$\delta_n = \frac{1}{(K-1)^2[\gamma_n * n + \int_{\gamma_n}^{1} \frac{dx}{F(x)}]} \tag{46}$$

Обозначим $\mathcal{I}_n = \int_{\gamma_n}^{1} \frac{dx}{F(x)}$, $\Psi(n)$ - произвольная растущая от n функция, $\lim\limits_{x\to\infty} \Psi(n) = \infty$

Теорема 1

Алгоритм $\mathcal{A}$ является асимптотически оптимальным при выполнении условий $\lim_{n\to\infty} \mathcal{I}_n = \infty$ и

$$\frac{b}{a} \leq \frac{1}{\Psi(n)} min\{\frac{1}{\gamma_n}, \frac{n}{\mathcal{I}_n}\} \tag{47}$$

Доказательство. Покажем, что при выполнении условий теоремы $\epsilon_n \to 0$ и $\delta_n \to 0$ с ростом n. Действительно, при $K > 1$ имеем:

$$\begin{aligned}\delta_n = \frac{1}{(\gamma_n n + \mathcal{I}_n)(K-1)^2} \leq \frac{1}{\mathcal{I}_n * (K-1)^2} \to 0 \\ \epsilon_n = K(\frac{b}{a} - 1)(\gamma_n + \frac{1}{n}\mathcal{I}_n) \leq K * \frac{b}{a}(\gamma_n + \frac{1}{n}\mathcal{I}_n) \leq \\ \leq \frac{K}{\Psi(n)} min(\frac{1}{\gamma_n}, \frac{n}{\mathcal{I}_n})(\gamma_n + \frac{1}{n}\mathcal{I}_n) = \\ = \frac{K}{\Psi(n)} min(1 + \frac{\mathcal{I}_n}{n\gamma_n}, 1 + \frac{n\gamma_n}{\mathcal{I}_n}) \leq \frac{2K}{\Psi(n)} \to 0\end{aligned} \tag{48}$$

при $n \to \infty$. Теорема 1 доказана.

Вероятностный анализ ИБГ для равномерного распределения

Определим условия асимптотической оптимальности алгоритма $\mathcal{A}$ для случая, когда веса ребер графа a_{ij} могут быть выбраны равновероятно из отрезка [a,b], a>0. В этом случае нормированная интегральная функция распределения имеет вид $F(x) = x,\ 0 \leq x \leq 1,\ \gamma_n = F(\frac{1}{n}) = \frac{1}{n}$ и

$$\mathcal{I}_n = \int_{\gamma_n}^{1} \frac{dx}{F(x)} = \int_{\frac{1}{n}}^{1} \frac{dx}{x} = ln(n) \tag{49}$$

Тогда из теоремы 1 непосредственно получаем результат, который может быть сформулирован как:

Теорема 2
Если элементы a_{ij} матрицы А принимают значения равновероятно из отрезка [a,b], то алгоритм $\mathcal{A}$ является асимптотически оптимальным при выполнении следующего условия:

$$\frac{b}{a} \leq \frac{n}{ln(n)} \frac{1}{\Psi(n)} \tag{50}$$

Представляет интерес оценить величины ϵ_n и δ_n, фигурирующие в соотношении (1).
Учитывая специфику равномерного распределения можно получить более точные оценки для этих величин по сравнению с общим случаем. Выведем условия асимптотической оптимальности алгоритма $\mathcal{A}$ в случае равномерного распределения, проведя в сокращенном виде вычисления оценок для $E(\mathcal{Z}_{\mathcal{A}})$ и $\mathcal{D}(\mathcal{Z}_{\mathcal{A}})$ и $P\{\mathcal{Z}_{\mathcal{A}} \leq (1+\epsilon_n)\mathcal{Z}\}$. Согласно (38)

$$E(l_k) = \int_0^1 (1-x)^{n-k} dx = \int_0^1 x^{n-k} dx = \frac{1}{n-k+1} \qquad (51)$$

$k = 1, 2, ...n-k$; $E(l_n) = E(l_n - 1) = \frac{1}{2}$
С учетом (39)

$$E(\mathcal{Z}_{\mathcal{A}}) = \sum_{k=1}^{n} [a + (b-a)E(l_k)] = na + (b-a)(\frac{1}{2} + \sum_{k=1}^{n-1} \frac{1}{n-k+1}) \leq$$

$$\leq na + (b-a)(\frac{1}{2} + ln(n)) = \mathcal{Z}^*_{\mathcal{A}} \qquad (52)$$

Оценим дисперсию $d_k = \int_0^1 (x - E(l_k))^2 d\Phi_k(x)$ случайной величины l_k с учетом того, что для равномерного распределения $\Phi_k(x)= 1 - (1-x)^{n-k}$ Используя (51), имеем

$$d_k = \int_0^1 x^2 d\Phi_k(x) - E(l_k)^2 = x^2\Phi_k(x)|_0^1 - 2\int_0^1 x\Phi_k(x)dx - E(l_k)^2 =$$
$$= 1 - 2\int_0^1 x[1-(1-x)^{n-k}]dx - E(l_k)^2 =$$
$$= \frac{2}{n-k+1} - \frac{2}{n-k+2} - \frac{1}{(n-k+1)^2},$$
$$k = 1, 2, ..., n-1;\ d_n = d_{n-1} = \frac{1}{12} \quad (53)$$

Отсюда с учетом определения дисперсии величины $\mathcal{Z}_{\mathcal{A}'}$, получим

$$
\begin{aligned}
\frac{1}{(b-a)^2} D(\mathcal{Z}_{\mathcal{A}'}) = \sum_{k=1}^{n} d_k = \frac{1}{12} + \sum_{k=1}^{n-1} \Big(\frac{2}{n-k+1} - \frac{2}{n-k+2} - \frac{1}{(n-k+1)^2} \\
= \frac{1}{12} - \frac{2}{n+1} + 1 - \sum_{k=1}^{n-1} \frac{1}{(n-k+1)^2} \leq \frac{13}{12} - \frac{2}{n+1} - \frac{1}{4} - \int_3^{n+1} \\
= \frac{13}{12} - \frac{2}{n+1} - \frac{1}{4} + \frac{1}{n+1} - \frac{1}{3} < 0.417 = \frac{\mathcal{D}^*}{(b -}
\end{aligned} \tag{54}
$$

Приведем оценку вероятности невыполнения соотношения (1) для случая равномерного распределения:

$$
\begin{aligned}
P\{\mathcal{Z}_{\mathcal{A}'} > (1+\epsilon_n)\mathcal{Z}\} \le P\{\mathcal{Z}_{\mathcal{A}'} > (1+\epsilon_n)na\} \le \\
\le P\{\mathcal{Z}_{\mathcal{A}'} + \mathcal{Z}^*_{\mathcal{A}'} - E(\mathcal{Z}_{\mathcal{A}'}) > (1+\epsilon_n)na\} = \\
= P\{\mathcal{Z}_{\mathcal{A}'} - E(\mathcal{Z}_{\mathcal{A}'}) > (1+\epsilon_n)na - \mathcal{Z}^*_{\mathcal{A}'}\} \le \\
\le P\{|\mathcal{Z}_{\mathcal{A}'} - E(\mathcal{Z}_{\mathcal{A}'})| > (1+\epsilon_n)na - \mathcal{Z}^*_{\mathcal{A}'}\} \le \\
\le \frac{D(\mathcal{Z}_{\mathcal{A}'})}{[(1+\epsilon_n)na + \mathcal{Z}^*_{\mathcal{A}'}]^2} \le \\
\le \frac{0.417(b-a)^2}{[1+\epsilon_n)na - (nu + (b-a)(\frac{1}{2} + ln(n)))]^2} = \\
= \frac{0.417}{[\frac{n\epsilon_n}{\frac{b}{a}-1} - ln(n) - \frac{1}{2}]^2}
\end{aligned}
\tag{55}
$$

Положим $\epsilon_n = \frac{c}{\Psi(n)}$, константа c>1, и пусть $\frac{a}{b} \le \frac{n}{ln(n)}\frac{1}{\Psi(n)}$. Тогда (55) может быть продолжено следующим образом:

$$\frac{0.417}{[\frac{n\frac{c}{\Psi(n)}}{(\frac{b}{a}-1)} - \frac{1}{2} - ln(n)]} \leq \frac{0.417}{[\frac{c\frac{b}{a}ln(n)}{(\frac{b}{a}-1)} - \frac{1}{2} - ln(n)]^2} = \frac{0.417}{[(c-1)ln(n) - \frac{1}{2}]^2} = \delta_n \tag{56}$$

$$\frac{b}{a} \leq \sqrt{\frac{n}{cln(n)}} \tag{57}$$

$$\epsilon_n \leq \sqrt{\frac{cln(n)}{n}} \tag{58}$$

Анализ соотношений **??**-58, полученных для равномерного распределения показывает, что уменьшение константы c "улучшает"оценки для $\frac{b}{a}$ и ϵ_n и "ухудшает"оценку вероятности $P\{\mathcal{Z}_{\mathcal{A}'} \leq (1+\epsilon_n)\mathcal{Z}\}$