

Глава 7.4. Избранные темы: Методы Монте-Карло

Методы Монте-Карло для марковских цепей (МСМС)

Одним из наиболее важных приложений цепей Маркова является моделирование случайных величин (принимающих значения в S). Пусть дана целевая мера $m \in \mathcal{P}(S)$, и мы заинтересованы в моделировании случайной величины X с распределением m , то есть $\mathbb{P}(X = x) = m_x$ для всех $x \in S$. Возможная стратегия состоит в том, чтобы запустить цепь Маркова (в идеале неприводимую, апериодическую) на S с инвариантной мерой m . При моделировании цепи в течение длительного времени T , распределение X_T будет сходиться **экспоненциально быстро** к целевой мере m . Конечно, важно (и крайне нетривиально) точно оценить скорость сходимости распределения X_T . Самое главное, при заданной целевой мере m , разработать такие вероятности перехода, которые гарантируют быструю сходимость и низкую вычислительную стоимость. Последнее свойство можно формализовать математически разными способами (например, избегание вычисления некоторых сложных величин, количество обращений к m_x за шаг, или просто то, что алгоритм моделирования использует ближайших соседей относительно некоторой естественной структуры на графе, где определена целевая мера m), ни один из которых не является полностью удовлетворительным.

Пример 0.1 (Имитация отжига). Предположим, мы хотим найти глобальный минимум функции стоимости $U : S \rightarrow \mathbb{R}$. Во многих приложениях (например, в комбинаторной оптимизации) пространство S очень велико, а у U много локальных минимумов, что делает алгоритмы локального или детерминированного поиска неэффективными.

Идея имитации отжига состоит в том, чтобы превратить эту задачу оптимизации в задачу моделирования (сэмплирования). Для параметра $\beta > 0$ (интерпретируемого как обратная температура) определим **меру Гиббса**:

$$m_x^\beta = \frac{1}{Z^\beta} e^{-\beta U(x)}, \quad Z^\beta = \sum_{y \in S} e^{-\beta U(y)}$$

При $\beta \rightarrow \infty$ (температура стремится к нулю) мера m^β сосредотачивается на множестве глобальных минимумов U . Если мы сможем получать выборку из m^β для очень больших β , мы найдем минимум с большой вероятностью (вероятность упустить его экспоненциально мала по β).

Однако прямое моделирование непрактично, например, из-за константы Z^β (вычисление Z^β подразумевает вычисление всех $U(x)$). Вместо этого мы запускаем цепь Маркова, разработанную так, чтобы m^β была её инвариантной мерой. Будем считать, что S имеет структуру графа со степенью, ограниченной D (то есть ни у одной точки нет более D соседей). Это мотивировано тем фактом, что, хотя $|S|$ во многих случаях велико, оно имеет структуру (связного, неориентированного) графа с *малой* степенью (см. Пример 0.2 ниже; или рассмотрите случай большого ящика $S = \{-N, \dots, N\}^d$ в \mathbb{Z}^d , тогда $|S| = (2N + 1)^d$, но $D = 2d$). Тогда мы возьмём

$$p_{x,y}^\beta = \begin{cases} e^{-\beta(U(y)-U(x))^+} / D & \text{если } x, y \text{ — соседи,} \\ 1 - \sum_{z \text{ сосед } x} e^{-\beta(U(z)-U(x))^+} / D & \text{если } y = x, \\ 0 & \text{иначе.} \end{cases} \quad (1)$$

где a^+ обозначает положительную часть вещественного числа a .

Этот выбор соответствует тому, что на данном шаге, когда $X_t = x$, мы выбираем с вероятностью $1/D$ заданного кандидата в соседи y для x . Если $U(y) \leq U(x)$, то $X_{t+1} = y$. Если $U(y) > U(x)$, то с вероятностью $e^{-\beta(U(y)-U(x))}$ мы полагаем $X_{t+1} = y$, а в противном случае не двигаемся, $X_{t+1} = x$. Идея состоит в том, что эта процедура, с одной стороны, благоприятствует переходам в сторону меньших значений U , а с другой стороны, позволяет (пусть и с малой вероятностью) «выбираться из ям» вокруг локальных минимумов U . Инвариантность m для этой случайной динамики является следствием более общего утверждения Утверждение 0.1.

На практике в так называемом алгоритме **имитации отжига** эту цепь моделируют с медленно растущим β (так что $\beta = \beta_t$) по мере увеличения времени t , что теоретически позволяет цепи избегать локальных минимумов и оседать в глобальном минимуме. Поскольку увеличение β означает снижение температуры со временем, название этого влиятельного метода происходит от **термической обработки** металлов.

Пример 0.2 (Задача коммивояжера). Рассмотрим N городов, помеченных $1, \dots, N$, и матрицу $(c_{i,j})_{1 \leq i, j \leq N}$, представляющую стоимость путешествия между городом i и городом j . Коммивояжёр должен посетить все N городов и вернуться домой, и он должен решить, каков оптимальный порядок посещения городов. Это означает, что мы ищем перестановку $\sigma \in \mathcal{S}_N$, которая минимизирует общую стоимость.

Здесь пространство состояний $S = \mathcal{S}_N$ — это множество всех перестановок N городов. Функция стоимости $U: S \rightarrow \mathbb{R}$ — это полная стоимость тура (мы понимаем суммы по модулю N , так что $N + 1 = 1$):

$$U(\sigma) = \sum_{i=1}^N c_{\sigma(i), \sigma(i+1)}$$

Чтобы применить метод из Пример 0.1, мы должны определить структуру графа на S . Попробуем разные стратегии.

- Две перестановки связаны, если одна может быть получена из другой путем **обмена** двух индексов. Если обозначить через $\pi^{i,j}\sigma$ транспозицию, меняющую местами элементы на позициях i и j , то соседями σ являются $D = \binom{N}{2}$ перестановок $(\pi^{i,j}\sigma)_{i < j}$.
- Две перестановки связаны, если одна может быть получена из другой путем обмена двух *соседних* индексов. Это означает, что соседями σ являются все перестановки, полученные из σ с помощью *смежной* транспозиции, а именно $D = N$ перестановок $(\pi^{i,i+1}\sigma)_i$.
- Если $c_{i,j} = c_{j,i}$ **симметрична**, две перестановки связаны, если одна может быть получена из другой путем *обращения* порядка посещения подсегмента индексов. Если обозначить через $\epsilon^{i,j}\sigma$ перестановку, полученную обращением пути между индексом i и j , то соседями σ являются $D = \binom{N}{2}$ перестановок $(\epsilon^{i,j}\sigma)_{i < j}$.

Теперь, как и в Пример 0.1, мы фиксируем большую константу $\beta > 0$ (которая управляет вероятностью принятия ходов «в гору»), и моделируем цепь Маркова с вероятностями перехода, заданными Уравнение 1. Решающее вычислительное преимущество локальных динамик а., б., с. состоит в том, что нам никогда не нужно вычислять полную стоимость $U(\sigma)$, а только разность $U(\cdot)$ между соседними перестановками. Она легко вычисляется (см. Рисунок 1) как

$$\delta U := \begin{cases} U(\pi^{i,j}\sigma) - U(\sigma) = (c_{\sigma(i-1), \sigma(j)} + c_{\sigma(j), \sigma(i+1)} + c_{\sigma(j-1), \sigma(i)} + c_{\sigma(i), \sigma(j+1)}) - (c_{\sigma(i-1), \sigma(i)} + c_{\sigma(i), \sigma(i+1)} + c_{\sigma(j-1), \sigma(j)}) \\ U(\pi^{i,i+1}\sigma) - U(\sigma) = (c_{\sigma(i-1), \sigma(i+1)} + c_{\sigma(i+1), \sigma(i)} + c_{\sigma(i), \sigma(i+2)}) - (c_{\sigma(i-1), \sigma(i)} + c_{\sigma(i), \sigma(i+1)} + c_{\sigma(i+1), \sigma(i+2)}) \\ U(\epsilon^{i,j}\sigma) - U(\sigma) = (c_{\sigma(i-1), \sigma(j)} + c_{\sigma(i), \sigma(j+1)}) - (c_{\sigma(i-1), \sigma(i)} + c_{\sigma(j), \sigma(j+1)}) \end{cases} \quad (2)$$

Таким образом, чтобы оценить Уравнение 1, нам никогда не нужно вычислять Z^β или даже U , только комбинацию из соответственно 8, 6, 4 слагаемых стоимости (для случаев а., б., с.) (независимо от N). Алгоритм тогда таков:

- Начните с произвольной перестановки, например, тождественной или случайной.

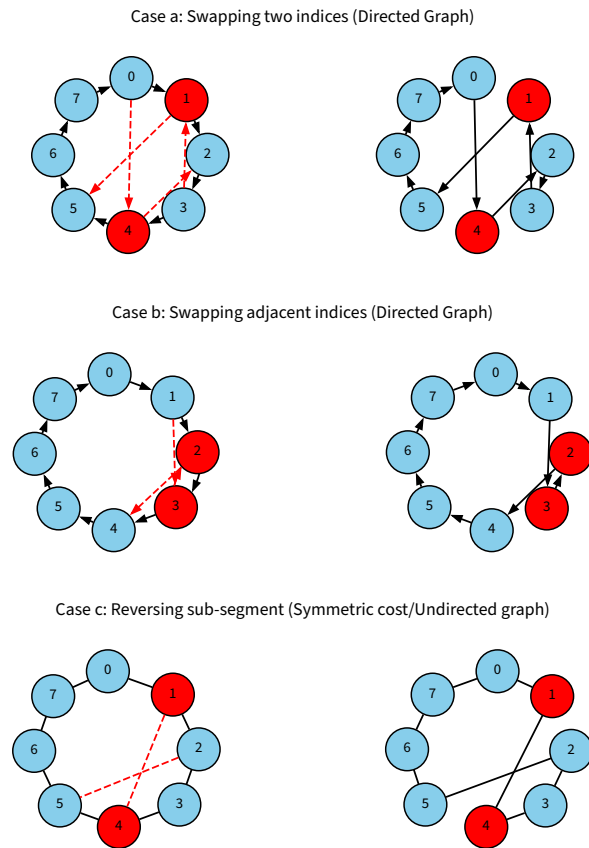


Рисунок 1: Различные одношаговые переходы для трёх примеров задачи Коммивояжёра.

2. Выберите два случайных индекса i, j (для динамики а.) или один случайный индекс i (для динамики б., так как $j = i + 1$ в этом случае).
3. Вычислите значение δU в Уравнение 2.
4. Если $\delta U \leq 0$, то выполните шаг (а именно, поменяйте местами i и j в случае а., поменяйте местами i и $i + 1$ в случае б., обратите путь между i и j в случае с.). Если $\delta U > 0$, то выполните шаг с вероятностью $e^{-\delta U}$, а в противном случае не делайте никаких изменений.
5. Повторяйте с пункта 2.

Конечно, критический вопрос заключается в том, сколько раз мы должны повторять шаги 2.-5. Это проблема, представляющая большой практический и теоретический интерес. Ответ критически зависит от N, β, U и допустимости возможных ошибок. Во многих ситуациях можно просто взять «достаточно большое T » и избежать математических сложностей.

В примерах а., б., с. выше, довольно ясно, чего ожидать:

- Случаи а. и с. относительно похожи, у каждой точки есть $N(N - 1)/2$ соседей, и им обычно требуется меньше шагов, чтобы обойти пространство состояний и «найти» точки минимума U . Обычно ожидается, что с. работает лучше, чем а. Динамика б. напротив допускает только N соседей и имеет более высокий шанс застрять в «локальных минимумах» U (а именно, чтобы найти путь меньшей стоимости, сначала нужно поменять местами несколько соседних индексов, при которых стоимость *возрастает*, но эти переходы маловероятны при больших β). В симуляциях мы видим, что б. действительно является алгоритмом, для которого U убывает медленнее всего с итерациями.
- Однако каждый отдельный шаг б. намного проще вычислительно, и если мы построим график $U(X_t)$ не как функцию количества шагов, а как функцию процессорного времени, затраченного на моделирование, разрыв между б. и другим алгоритмом значительно сократится.

- В конечном счёте, мы видим, что увеличение β по ходу моделирования является лучшей стратегией, и, безусловно, для работы b. нам просто нужно настраивать $\beta \equiv \beta_t$ так, чтобы она росла, но медленно по сравнению с a. и b.

В конечном счёте, метод сводится к следующему подходу, называемому МСМС (Markov Chain Monte Carlo — методы Монте-Карло для марковских цепей). При заданной $m \in \mathcal{P}(S)$, найти некоторые «вычислительно оптимальные» вероятности перехода $(p_{x,y}^t)_{t \geq 0, x, y \in S}$, такие что $\sum_y m_y p_{y,x}^t = m_x$, для всех $t \in \mathbb{N}, x \in S$. Мы сосредоточимся на однородном по времени случае, когда $p_{x,y}^t \equiv p_{x,y}$ не зависит от t .

Приведем несколько примеров, предполагая (без потери общности), что $m_x > 0$ для всех $x \in S$.

1. $p_{x,y} = m_y$, **тривиальное точное моделирование**, дающее цепь Маркова, которая является просто последовательностью н.о.р. величин. В этом случае распределение X_t сходится к m всего за один шаг.
2. $p_{x,y} = \varepsilon \sqrt{m_y/m_x} \mathbf{1}_{x \neq y} + (1 - \varepsilon \sum_{z \neq x} \sqrt{m_z/m_x}) \mathbf{1}_{x=y}$, известный как **глобальная тепловая баня** (global heat bath). Здесь $0 < \varepsilon \leq \inf_{x,y} \sqrt{m_x/m_y}$, так что $p_{x,y}$ действительно является вероятностью перехода с **обратимой**, и, следовательно, инвариантной мерой m .
3. $p_{x,y} = A_{x,y} \sqrt{m_y/m_x}$, обобщая предыдущие примеры. Здесь $A_{x,y}$ — *достаточно малая* симметричная матрица, в том смысле, что

$$A_{x,y} = A_{y,x} \geq 0, \quad x, y \in S, \quad A_{x,x} \sqrt{m_x} = 1 - \sum_{z \neq x} S_{x,z} \sqrt{m_z}, \quad x \in S \quad (3)$$

Ясно, что $m_x p_{x,y} = A_{x,y} \sqrt{m_x m_y} = m_y p_{y,x}$, так что m является **обратимой**, и, следовательно, инвариантной.

В типичных приложениях S **велико**, а m_x — какая-то *сложная* мера. Первые два примера выше крайне непрактичны. В первом случае мы просто моделируем m детерминированно. Однако мгновенная сходимость подразумевает очень высокую вычислительную стоимость единственного необходимого шага. В глобальной тепловой бане на каждом шаге нам нужно знать отношения $\sqrt{m_y/m_x}$ для всех $x, y \in S$, и несмотря на то, что это, как правило, вычислить проще, чем m_x (см. Пример 0.2), это всё же требует доступа к $|S|^2$ отношениям значений инвариантной меры. Третий общий случай более интересен (и будет дополнительно обобщен ниже), но пока мало что говорит нам о хорошем выборе $A_{x,y}$.

Динамика Метрополиса — Гастингса

Задача становится намного интереснее как математически, так и практически, если предположить, что задана некоторая дополнительная структура. В частности, предположим, что даны некоторые *исходные* вероятности перехода $q_{x,y}$. Осторожно: мы не предполагаем здесь, что m инвариантна для $q_{x,y}$; скорее, $q_{x,y}$ следует воспринимать как вытекающие из некоторой особенности, например геометрических свойств, рассматриваемой задачи. Если S имеет структуру графа, например, $q_{x,y}$ могут быть вероятностями перехода **простого случайного блуждания** на S .

Утверждение 0.1 (Динамика Метрополиса — Гастингса). Пусть S конечно, и предположим, что даны:

- **Целевая мера** $m \in \mathcal{P}(S)$.
- **Ядро предложения** (proposal kernel) $Q = (q_{x,y})_{x,y}$, которое является марковским переходным ядром.
- **Кососимметричное вихревое ядро** $R = (r_{x,y})_{x,y}$, такое что $r_{x,y} = -r_{y,x}$ и $\sum_y r_{x,y} = 0$.

Предположим, что $r_{x,y} \geq -m_y q_{y,x}$ для всех $x, y \in S$ (в частности, так как R кососимметрично, $-m_y q_{y,x} \leq r_{x,y} \leq m_x q_{x,y}$). Тогда

$$p_{x,y} := \begin{cases} \min(q_{x,y}, \frac{r_{x,y} + m_y q_{y,x}}{m_x}) & \text{если } y \neq x \\ 1 - \sum_{z \neq x} \min(q_{x,z}, \frac{r_{x,z} + m_z q_{z,x}}{m_x}) & \text{если } y = x \end{cases} \quad (4)$$

определяет допустимые марковские вероятности перехода, называемые (обобщённым) **ядром Метрополиса — Гастингса**, которое допускает m в качестве инвариантной вероятности.

Предположим более того, что $q_{x,y}$ неприводимо и что $q_{x,y} > 0$, когда $q_{y,x} > 0$. Тогда $p_{x,y}$ неприводимо, и, следовательно, m является единственной инвариантной вероятностью.

Доказательство. Свойство нормировки $\sum_y p_{x,y} = 1$ следует непосредственно из определения $p_{x,x}$. Далее проверим, что $p_{x,y} \geq 0$. Это как раз условие $r_{x,y} \geq -m_y q_{y,x}$ для внедиагональных членов. В то же время $p_{x,x} \geq 0$, так как мы можем ограничить сумму с отрицательным знаком величиной $\sum_{z \neq x} q_{x,z} = 1 - q_{x,x} \leq 1$.

Далее проверим, что m инвариантна. Действительно,

$$\begin{aligned} \sum_{y \neq x} m_x p_{x,y} &= \sum_{y \neq x} \min(m_x q_{x,y}, r_{x,y} + m_y q_{y,x}) = \sum_{y \neq x} \min(m_x q_{x,y} - r_{x,y}, m_y q_{y,x}) + \sum_{y \neq x} r_{x,y} \\ &= \sum_{y \neq x} \min(m_x q_{x,y} + r_{y,x}, m_y q_{y,x}) = \sum_{y \neq x} m_y p_{y,x} \end{aligned}$$

где при переходе от первой ко второй строке мы использовали $r_{x,y} = -r_{y,x}$ и $\sum_{y \neq x} r_{x,y} = 0$ (так как $r_{x,x} = 0$). \square

Алгоритмы Метрополиса — Гастингса работают очень хорошо во многих ситуациях, однако они имеют некоторые слабые стороны:

- Если наша задача состоит в том, чтобы сойтись к инвариантной мере как можно быстрее, наличие шанса $p_{x,x} > 0$ не двигаться вообще вряд ли оптимально.
- При увеличении β наличие локальных минимумов U может резко увеличить время сходимости цепи Маркова.

Упражнение 0.1. Проверьте, что если $r = 0$, то m обратима для вероятностей перехода $p_{x,y}$, определённых в Уравнение 4.

Упражнение 0.2. Проверьте, что динамика Метрополиса — Гастингса включает в себя вышеупомянутые примеры: тривиальное точное моделирование, глобальную тепловую баню и симметричную динамику Уравнение 3. (Покажите, что для конкретного выбора $r_{x,y}, q_{x,y}$ получаются эти динамики).

Динамики Глаубера и Кавасаки

Динамика Глаубера — это частный случай МСМС, часто используемый в статистической физике для спиновых систем (например, модели Изинга), соответствующий «одноузловой тепловой бане». В частности, рассмотрим случай, когда S — это пространство-произведение, скажем $S = \{-1, +1\}^V$, для некоторого конечного графа V . Мы интерпретируем S как *спиновое пространство*: каждая точка в V имеет признак, скажем спин, который в данный момент времени может быть либо $+1$, либо -1 . Более того, V сам по себе является графом, а именно, у нас есть понятие «соседей» данной вершины $v \in V$. Целевая мера — это мера Гиббса $m(\sigma) \propto e^{-\beta H(\sigma)}$, где Гамильтониан — это функция $H: S \rightarrow \mathbb{R}$, например $H(\sigma) = -\sum_{(u,v) \in E} J_{u,v} \sigma_u \sigma_v$. На каждом шаге динамика происходит следующим образом:

1. Выбрать вершину $v \in V$ равномерно случайно.
2. Заменить спин σ_v новым значением $\sigma'_v \in \{-1, +1\}$, выбранным из условного распределения спина в v при условии конфигурации его соседей.

Вероятность перехода для переворота спина в узле v (изменения σ на σ^v) задаётся формулой:

$$p_{\sigma, \sigma^v} = \frac{1}{|V|} \frac{e^{-\beta H(\sigma^v)}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma^v)}} = \frac{1}{|V|} \frac{1}{1 + e^{\beta(H(\sigma^v) - H(\sigma))}}$$

Эта динамика обратима относительно меры Гиббса. Ключевая особенность состоит в том, что разность энергий $H(\sigma^v) - H(\sigma)$ зависит *только* от соседей v , что делает обновление очень дешёвым вычислительно ($O(1)$, если степень графа ограничена).

В то время как динамика Глаубера позволяет изменяться полной «магничности» (сумме спинов), **динамика Кавасаки** предназначена для получения выборки из меры Гиббса при условии фиксированного числа спинов $+1$ (канонический ансамбль). Например, мы можем думать в этом случае, что $S = \{0, 1\}^V$, где 0 представляет пустое место, а 1 — место, занятое частицей. Нам нужна динамика, сохраняющая общее число частиц. Например:

1. Выбрать пару вершин (u, v) , соединённых ребром в графе.
2. Предложить **поменять местами** значения σ в u и v .
3. Принять обмен с вероятностью, определённой правилом Метрополиса:

$$p_{\sigma, \sigma^{u,v}} = \min(1, e^{-\beta(H(\sigma^{u,v}) - H(\sigma))})$$

где $\sigma^{u,v}$ — это конфигурация σ , в которой значения в u и v поменяны местами.

Если $\sigma_u = \sigma_v$, обмен ничего не меняет, и энергия постоянна. Если они различаются, частица перемещается. Поскольку частицы только перемещаются и никогда не создаются и не уничтожаются, общее количество спинов $+1$ сохраняется на протяжении всей эволюции цепи.

Точное моделирование

Наконец, мы обсудим несколько иную ситуацию, где мы хотим получить выборку из *неизвестной* инвариантной меры данной цепи Маркова, так что вероятности перехода в этом случае *даны*. Мы предполагаем, что вероятности перехода удовлетворяют следующему условию

$$Z := \sum_y \inf_{x \in S} p_{x,y} > 0 \quad (5)$$

Другими словами, существует хотя бы одна точка, в которую цепь Маркова может перепрыгнуть с равномерно положительной вероятностью независимо от начальной точки x . Хотя это может показаться очень ограничительным, даже если $P = (p_{x,y})_{x,y}$ не удовлетворяет Уравнение 5, возможно, некоторая степень P^t будет удовлетворять. И в этом случае мы можем просто применить ту же технику к вероятности перехода за t шагов. В частности, это сработает для любой апериодической цепи на конечном или счётном пространстве состояний.

Примечание. Предполагая Уравнение 5, мы можем записать

$$p_{x,y} = (1 - \varepsilon)q_{x,y} + \varepsilon\nu_y \quad (6)$$

для некоторых $\varepsilon > 0$, $\nu \in \mathcal{P}(S)$ и марковских вероятностей перехода $Q = (q_{x,y})_{x,y}$. Например, мы можем взять $\varepsilon = Z$ и

$$q_{x,y} = (p_{x,y} - \varepsilon\nu_y)/(1 - \varepsilon)$$

Заметим, что если $p_{x,y}$ неприводима (или, более общо, рассматривая только сообщающийся класс, содержащий носитель ν), то P индуцирует **положительно возвратную** цепь Маркова (так как на каждом шаге мы имеем вероятность $\varepsilon\nu_y > 0$ вернуться в некоторую точку y в носителе ν). В частности, цепь Маркова допускает **единственное инвариантное распределение** $m \in \mathcal{P}(S)$.

Утверждение 0.2 (Алгоритм точного моделирования). *Предположим, неприводимая цепь Маркова на конечном или счётном пространстве состояний S , с вероятностями перехода $P = (p_{x,y})$, удовлетворяет Уравнение 5, и пусть m — её инвариантная вероятность. Для любого разложения $p_{x,y} = (1 - \varepsilon)q_{x,y} + \varepsilon\nu_y$ как в Уравнение 6 выполняется следующее.*

Пусть τ — геометрическая случайная величина с параметром ε , то есть $\mathbb{P}(\tau = k) = \varepsilon(1 - \varepsilon)^k$ для $k \in \mathbb{N}$. Пусть $\mathbf{Y} := (Y_t)$ — цепь Маркова с вероятностями перехода $Q = (q_{x,y})$ и начальным распределением ν . Тогда Y_τ имеет распределение m .

Доказательство. Пусть μ — распределение Y_τ , тогда для $y \in S$

$$\mu_y = \sum_{k=0}^{\infty} \mathbb{P}(Y_\tau = y | \tau = k) \mathbb{P}(\tau = k) = \sum_{k=0}^{\infty} \varepsilon(1 - \varepsilon)^k \mathbb{P}(Y_k = y)$$

Следовательно, $\mu = \varepsilon \sum_k (1 - \varepsilon)^k \nu Q^k$. Нам нужно проверить $\mu = \mu P$, так как мы уже знаем, что инвариантная мера единственна, и единственным решением этого уравнения является $\mu = m$.

В операторной записи $P = \varepsilon R + (1 - \varepsilon)Q$, где $R_{x,y} = \nu_y$ для всех $y \in S$. Заметим, что $\lambda R = \nu$ для всех $\lambda \in \mathcal{P}(S)$. Следовательно,

$$\begin{aligned} \mu P &= \varepsilon \mu R + (1 - \varepsilon) \mu Q = \varepsilon \nu + (1 - \varepsilon) \left(\varepsilon \sum_{k=0}^{\infty} (1 - \varepsilon)^k \nu Q^k \right) Q = \varepsilon \left(\nu + \sum_{k=0}^{\infty} (1 - \varepsilon)^{k+1} \nu Q^{k+1} \right) \\ &= \varepsilon \left(\nu(1 - \varepsilon)^0 \nu Q^0 + \sum_{k=1}^{\infty} (1 - \varepsilon)^k \nu Q^k \right) = \mu \end{aligned}$$

□

Это означает, что если ν и $q_{x,y}$ доступны для вычислений, у нас есть очень практичный способ получить выборку из m :

1. Смоделируйте геометрическую случайную величину τ с параметром успеха ε (например, подсчитайте количество неудач до первого успеха при подбрасывании монеты с ε).
2. Смоделируйте точку $x_0 \in S$ с распределением ν .
3. Смоделируйте τ шагов цепи Маркова, начинающейся в x_0 , с вероятностями перехода $(q_{x,y})$.
4. Положение цепи Маркова после τ шагов является выборкой из m .

Этот подход является точным (perfect), поскольку мы не утверждаем никакой сходимости в долгосрочной асимптотике. Это просто точное сэмплирование, хотя количество итераций, необходимых для моделирования (Y_t) , случайно и (с малой вероятностью) может быть очень большим.