

Глава 6: Спектральные свойства цепей Маркова

В этой главе мы сосредоточимся на цепях Маркова с конечным числом состояний, так что линейный оператор P может быть просто представлен матрицей, и мы можем использовать некоторые известные спектральные свойства конечномерных линейных операторов. Мы видели, что $P\mathbf{1} = \mathbf{1}$, то есть собственное значение 1 принадлежит спектру P . И действительно, P также допускает левый собственный вектор, инвариантная вероятность m такая, что $mP = m$.

Теорема 0.1 (Теорема Перрона-Фробениуса для стохастических матриц). Пусть P — матрица переходных вероятностей цепи Маркова с конечным пространством состояний S . Тогда:

- Постоянная функция $\mathbf{1}$ (рассматриваемая как вектор со всеми элементами, равными 1) является собственным вектором с собственным значением $\lambda_0 = 1$. Любая инвариантная вероятность является левым собственным вектором матрицы P с положительными компонентами, ассоциированным с тем же собственным значением.
- Любое собственное значение $\lambda \in \text{Spec}(P)$ удовлетворяет $|\lambda| \leq 1$.

Предположим теперь, что цепь неприводима.

- Собственное значение $\lambda_0 = 1$ простое, и поэтому соответствующий левый собственный вектор является с точностью до постоянного множителя единственной инвариантной вероятностной мерой m . Все компоненты m строго положительны.

Предположим теперь, что цепь также аperiodична.

- Все остальные собственные значения λ матрицы P удовлетворяют $|\lambda| < 1$.

Доказательство.

- следует непосредственно из $\sum_y p_{x,y} = 1$ для всех $x \in S$, что означает $P\mathbf{1} = \mathbf{1}$. Аналогично $mP = m$ для любой инвариантной вероятности m .
- Пусть $\lambda \in \text{Spec}(P)$ — произвольное собственное значение и v — соответствующий собственный вектор (возможно, с комплексными компонентами), $Pv = \lambda v$. Для $t \geq 0$ имеем $P^t v = \lambda^t v$ и поэтому (мы рассматриваем произвольное t для дальнейшего использования этого неравенства, иначе мы могли бы просто взять $t = 1$ здесь)

$$|\lambda^t v_x| = |(P^t v)_x| = \left| \sum_{y \in S} p_{x,y}^{(t)} v_y \right| \leq \sum_{y \in S} p_{x,y}^{(t)} \sup_z |v_z| = \sup_z |v_z| \quad (1)$$

Оптимизируя по x получаем $|\lambda^t| \sup_x |v_x| \leq \sup_z |v_z|$, то есть $|\lambda| \leq 1$, так как $v \neq 0$. Иными словами, спектральный радиус P равен 1.

- с. Зафиксируем собственный вектор v , ассоциированный с собственным значением $\lambda_0 = 1$. Выберем x такое, что $|v_x| = \sup_z |v_z|$, и нормируем собственный вектор так, чтобы $v_x = 1$. Тогда имеем для каждого $x \in S$

$$1 = v_x = (P^t v)_x = \sum_{y \in S} p_{x,y}^{(t)} v_y \leq \sup_z |v_z| = 1 \quad (2)$$

Таким образом, все неравенства обращаются в равенства, поэтому $v_y = v_x = 1$ всякий раз, когда $p_{x,y}^{(t)} > 0$. Но цепь неприводима, поэтому для каждого $y \in S$ существует некоторое $t \geq 0$, для которого это выполняется, и значит $v = \mathbf{1}$.

- d. Зафиксируем теперь собственный вектор v , ассоциированный с собственным значением $|\lambda| = 1$. Как и в п. с., мы можем нормировать v так, чтобы $v_x = \sup_z |v_z| = 1$. Мы хотим доказать (предполагая неприводимость и апериодичность), что $v = \mathbf{1}$, в частности $\lambda = 1$. Последнее неравенство в п. b. дает, для $|\lambda| = 1$, $1 \leq \sup_z |v_z| = 1$, поэтому все неравенства в Уравнение 1 становятся равенствами. Но тогда, всякий раз когда $p_{x,y}^{(t)}, p_{x,z}^{(t)} > 0$, имеем $v_y = v_x$. Поскольку цепь апериодична и неприводима, для достаточно большого t все переходные вероятности строго положительны, и поэтому $v_y = v_x = 1$ для всех $y \in S$. □

Теорема Перрона-Фробениуса предоставляет мощный альтернативный способ понять сходимость к стационарному распределению. Если P диагонализуема, мы можем записать её спектральное разложение. Для любого начального распределения π распределение в момент времени t равно πP^t . Теорема подразумевает, что при $t \rightarrow \infty$ слагаемые, ассоциированные с собственными значениями $|\lambda| < 1$, экспоненциально быстро стремятся к нулю, оставляя только проекцию на стационарное распределение.

Экспоненциальная сходимость для цепей с конечным числом состояний

Мы видели, что для неприводимых положительно-возвратных апериодических цепей закон X_t сходится к инвариантной мере при $t \rightarrow \infty$. В частности, это имеет место для неприводимых конечных (следовательно, положительно-возвратных) апериодических цепей. Для конечных цепей, однако, мы можем получить более точный результат о сходимости. Следующее предложение показывает, что если λ_1 — собственное значение (отличное от $\lambda_0 = 1$) матрицы P с наибольшим абсолютным значением, то

$$\sum_y |\mathbb{P}_x(X_t = y) - m_y| \simeq |\lambda_1|^{t(1+o(1))}$$

то есть, чем меньше $|\lambda_1|$, тем быстрее закон X_t сходится к m .

Теорема 0.2. Пусть P — матрица переходных вероятностей неприводимой апериодической цепи Маркова с конечным пространством состояний S (содержащим по крайней мере два элемента) с инвариантной мерой m , и пусть $\mu_t^x := (\delta_x P^t)$ — закон X_t когда $X_0 = x$. Пусть $\varrho := -\sup_{\lambda \in \text{Spec}(P) \setminus \{1\}} \log |\lambda| > 0$. Тогда

$$\sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mu_t^x - m\|_{TV} := \sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \left(\sum_y |p_{x,y}^{(t)} - m_y| \right) = \varrho$$

Доказательство. Поскольку $\lambda_0 = 1$ простое, мы можем записать $P = P_0 + Q$, где P_0 — оператор проектирования ранга 1, удовлетворяющий $\mu P = m$ для всех $\mu \in \mathcal{P}(S)$, и $P_0 Q = Q P_0 = 0$. Иными словами, мы можем записать P в жордановой форме как

$$P = V \begin{pmatrix} 1 & 0 \\ 0 & J \end{pmatrix} V^{-1}, \quad P_0 = V \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} V^{-1}, \quad Q = V \begin{pmatrix} 0 & 0 \\ 0 & J \end{pmatrix} V^{-1}.$$

и ясно, что P_0 и Q удовлетворяют упомянутым свойствам. В частности, $P^t = P_0 + Q^t$ для $t \geq 1$ и $|Q_{x,y}^t| \leq C t^{|S|} e^{-t\varrho}$, так как $e^{-\varrho}$ — модуль наибольшего собственного значения Q . Получаем $\delta_x P^t = \delta_x P_0 + \delta_x Q^t = m + \delta_x Q^t$. Таким образом,

$$\sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mu_t^x - m\|_{TV} \leq \sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log (|S| \sup_y |Q_{x,y}^t|) \leq \sup_{x,y} \lim_{t \rightarrow \infty} \frac{1}{t} \log (|S| C t^{|S|} e^{t\varrho}) = \varrho \quad (3)$$

С другой стороны, множество $(\delta_x)_{x \in S}$ представляет $|S|$ линейно независимых векторов. Таким образом (поскольку мы предположили $|S| \geq 2$) существует по крайней мере одно $x \in S$, такое что δ_x имеет ненулевую компоненту вдоль собственного вектора, ассоциированного с собственным значением λ_1 максимального модуля. Нетрудно проверить, что для такого x равенство достигается в Уравнение 3, поскольку $|\delta_x Q^t|$ ведёт себя как $c|\lambda_1|^t$ в этом случае. \square

Дополнительные Задачи

Упражнение 0.1. Рассмотрите положительно-возвратную аperiodическую цепь. Если $f: S \rightarrow \mathbb{R}$ такова, что $m(f) = 0$, то $m(Pf) = 0$. Иными словами, ограничение \hat{P} оператора P на функции с нулевым m -средним корректно определено. Повторите предыдущее доказательство, используя напрямую оператор \hat{P} .

Упражнение 0.2. Найдите пример цепи Маркова, для которой существует $x \in S$, такое что $\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mu_t^x - m\|_{TV} < \varrho$. Подсказка: достаточно рассмотреть пространство с 3 точками.

Упражнение 0.3. Пусть X — аperiodическая цепь Маркова с пространством состояний S . Пусть $\xi \in S$ — поглощающая точка, и пусть $S' = S \setminus \{\xi\}$. Предположим, что $x \leftrightarrow y$ для всех $x, y \in S'$, $x \rightarrow \xi$ для всех $x \in S'$, но $p_{\xi, \xi} = 1$, так что $\xi \nleftrightarrow x$ для $x \in S'$. Например, ξ представляет состояние, в котором случайный алгоритм завершается.

Пусть $\tau \equiv \tau_\xi$ — время достижения ξ , и пусть $Q := (p_{x,y})_{x,y \in S'}$. Докажите, что существует вероятностная мера $m \in \mathcal{P}(S')$, такая что для всех $x, y \in S'$

$$\lim_t \mathbb{P}_x(X_t = y | \tau > t) = m_y$$

Упражнение 0.4. Пусть P — марковский оператор неприводимой цепи Маркова на конечном пространстве состояний S . Пусть m — её инвариантная мера, а P^\dagger — сопряжённый оператор к P в $L^2(m)$.

- Проверьте, что если P, Q — марковские операторы, то PQ и $\alpha P + (1 - \alpha)Q$ также являются марковскими операторами. В частности, PP^\dagger и P^2 — марковские операторы.
- Найдите инвариантную меру для PP^\dagger и P^2 . Приведите пример, где P неприводима, но PP^\dagger и P^2 не являются таковыми.
- Приведите пример, где P неприводима, но P^2 и PP^\dagger не являются неприводимыми.
- Докажите, что если P неприводима и аperiodична, то P^2 неприводима и аperiodична.
- Докажите, PP^\dagger аperiodична и каждый класс общения закрыт.
- Предположим, что P неприводима и аperiodична. Пусть X_t и Y_t — цепи Маркова с марковскими операторами P^2 и PP^\dagger соответственно; докажите, что X_t сходится к своему «равновесию» быстрее, чем Y_t , в следующем смысле. Пусть μ_t^x (соответственно ν_t^x) — закон распределения X_t при условии $X_0 = x$ (соответственно Y_t при $Y_0 = x$). Докажите, что

$$\sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mu_t^x - m\|_{TV} \leq \sup_x \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\nu_t^x - m\|_{TV}$$

💡 Решение

- a. Пусть $R = PQ$, тогда $r_{x,z} = \sum_y p_{x,y} q_{y,z}$. Меняя порядок суммирования, имеем $r_{x,z} \geq 0$ и $\sum_z r_{x,z} = \sum_y p_{x,y} \sum_z q_{y,z} = \sum_y p_{x,y} = 1$. Что касается выпуклой комбинации, доказательство следует непосредственно из линейности.
- b. Инвариантной мерой всех этих операторов является m . В общем случае, если P, Q имеют одну и ту же инвариантную меру, то $m((PQ)f) = m(P(Qf)) = m(Qf) = m(f)$. Таким образом, m инвариантна для PQ .
- c. Мы можем взять единственный цикл $S = \mathbb{Z}_{2k}$ с чётным числом точек, P соответствует движению по часовой стрелке с вероятностью 1, $p_{x,x+1} = 1$ (суммы понимаются $(\text{mod } 2k)$). Тогда P неприводима, так как существует путь, соединяющий любые две точки. PP^\dagger соответствует выполнению одного шага по часовой стрелке и одного против часовой стрелки, то есть $PP^\dagger = I$, блуждание просто не движется. P^2 соответствует выполнению двух шагов по часовой стрелке $p_{x,x+2}^{(2)} = 1$. Таким образом, не существует пути, соединяющего точки с разной чётностью.
- d. Для неприводимой апериодической цепи на конечном пространстве состояний существует достаточно большое T , такое что $p_{x,y}^{(t)} > 0$ для всех $t > T$ и $x, y \in S$. В частности, то же самое верно, если заменить T на $T' = \lceil T/2 \rceil$, для элементов P^2 .
- e. Обозначим через $(q_{x,y})$ элементы $Q = PP^\dagger$. $q_{x,z} > 0$ тогда и только тогда, когда существует $y \in S$, такое что $p_{x,y}, p_{z,y} > 0$. В частности, $q_{x,z} > 0$ тогда и только тогда, когда $q_{z,x} > 0$ и $q_{x,x} > 0$ для всех $x \in S$ (так как каждая точка x имеет по крайней мере одного *преемника* y_x , такого что $p_{x,y_x} > 0$). В частности, каждый класс сообщающихся состояний Q замкнут.
- f. Согласно [теореме об экспоненциальной сходимости](#), нам нужно проверить, что $|\lambda_1(P^2)| \leq |\lambda_1(PP^\dagger)|$, где $\lambda_1(Q)$ — наибольшее (по модулю) собственное число неприводимого марковского оператора Q , за исключением $\lambda_0 = 1$; или, что эквивалентно, наибольшее собственное число Q , ограниченного на функции f с $m(f) = 0$, где m — инвариантная мера Q .

Для P^2 имеем $\lambda_1(P^2) = \lambda_1(P)^2 = \overline{\lambda_1(P^\dagger)^2}$. Возьмём f_1 — ассоциированную (комплексную) собственную функцию (с $m(f_1) = 0$) оператора P^\dagger , так что $P^\dagger f_1 = \overline{\lambda_1(P^\dagger)} f_1$. Тогда, так как PP^\dagger симметричен, $\lambda_1(PP^\dagger) > 0$, и, обозначая через $\langle \cdot, \cdot \rangle$ эрмитово скалярное произведение в $L^2(m)$, имеем

$$|\lambda_1(PP^\dagger)| = \sup_{f: m(f)=0} \frac{\langle f, PP^\dagger f \rangle}{m(|f|^2)} = \sup_{f: m(f)=0} \frac{\langle P^\dagger f, P^\dagger f \rangle}{m(|f|^2)} \geq \frac{\langle P^\dagger f_1, P^\dagger f_1 \rangle}{m(|f_1|^2)} = \overline{\lambda_1(P^\dagger)} \lambda_1(P^\dagger) = |\lambda_1(P^2)|$$

i Абстракция

Если S счётно (или более общо, польское пространство), предыдущая теорема может оказаться неверной. Это связано с тем, что собственные значения могут накапливаться вокруг одной точки, или, иначе говоря, одно собственное значение может иметь бесконечную кратность. Важный класс операторов, для которых это не может произойти, — **компактные операторы**.

Рассмотрим неприводимую положительно-возвратную аperiodическую цепь, где S — счётное пространство. Тогда она допускает единственную инвариантную вероятность m , и P является **сжатием** на $P: L^2(m) \rightarrow L^2(m)$. В частности, он отображает замкнутый шар $B_1 := \{f : \|f\|_{L^2(m)} \leq 1\}$ в себя. P называется **компактным** (линейным оператором), если он отображает замкнутый шар B_1 в **компактное** подмножество $L^2(m)$ (это всегда имеет место, если S конечно, поскольку B_1 сам по себе компактен в этом случае). Если P компактен, мы также имеем $\varrho := -\sup_{\lambda \in \text{Spec}(P) \setminus \{1\}} \log |\lambda| > 0$, и Теорема 0.2 переносится на случай счётного пространства состояний с компактным оператором P (поскольку спектр компактных операторов состоит из счётного множества собственных значений без точек накопления, кроме 0, каждое собственное значение имеет конечную кратность, поэтому доказательство в основном то же самое).

Расширения этого результата возможны; в частности, если P **квазикомпактен**, утверждение также остаётся верным. Это особенно полезно в динамических системах. Опишем быстро этот случай, предполагая S счётным, чтобы не обсуждать вопросы измеримости.

Если у нас есть отображение $\phi: S \rightarrow S$, мы можем рассматривать $(X_0, X_1 = \phi(X_0), X_2 = \phi(X_1), \dots)$ как цепь Маркова, хотя довольно специфическую: вся случайность исходит от X_0 , поскольку ϕ просто детерминировано. Здесь переходная вероятность просто

$$p_{x,y} = \delta_{\phi(x)}(y) = \begin{cases} 1 & \text{если } y = \phi(x) \\ 0 & \text{иначе} \end{cases}$$

Если m — инвариантная вероятность, мы знаем, что сопряжённый оператор P^\dagger также является оператором перехода цепи Маркова. Классический инструмент для изучения перемешивания динамических систем (например, экспоненциальной сходимости к инвариантной вероятности m) действительно — квазикомпактность P^\dagger .