

Г. А. Амбарцумян

К энтропии цепей Маркова

В математической и технической литературе встречаются различные определения энтропии цепей Маркова. В первой части настоящей работы устанавливается связь между этими определениями для сложных цепей. Во второй части—обобщается предельная теорема Хинчина [1] для простых цепей, на сложные цепи Маркова любого порядка.

1. Различные определения энтропии для сложных цепей Маркова и связь между ними

1. Рассмотрим опыты, образующие стационарную цепь Маркова порядка μ с конечным числом возможных состояний при каждом опыте:

$$A_1, A_2, \dots, A_n$$

соответственно с вероятностями

$$p(1), p(2), \dots, p(n),$$

где

$$p(\alpha) = P(A_\alpha) \quad (\alpha = 1, 2, \dots, n)$$

есть вероятность состояния A_α и

$$\sum_{\alpha=1}^n p(\alpha) = 1.$$

Обозначим

$$P_{\alpha_1 \alpha_2 \dots \alpha_m}(A_{\alpha_{m+1}}) = p_{\alpha_1 \alpha_2 \dots \alpha_m}(\alpha_{m+1})$$

вероятность состояния $A_{\alpha_{m+1}}$, при $(m+1)$ -ом опыте, если известно, что при предыдущих m опытах имели место состояния: $A_{\alpha_1}, A_{\alpha_2}, \dots, A_{\alpha_m}$.

Для цепи порядка μ будут иметь место равенства:

$$p_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}(\alpha_{\mu+2}) = p_{\alpha_2 \alpha_3 \dots \alpha_{\mu+1}}(\alpha_{\mu+2})$$

$$p_{\alpha_1 \alpha_2 \dots \alpha_{\mu+2}}(\alpha_{\mu+3}) = p_{\alpha_2 \alpha_3 \dots \alpha_{\mu+2}}(\alpha_{\mu+3})$$

и вообще при $m \geq \mu$:

$$p_{\alpha_1 \alpha_2 \dots \alpha_m}(\alpha_{m+1}) = p_{\alpha_{m-\mu+1} \alpha_{m-\mu+2} \dots \alpha_m}(\alpha_{m+1}).$$

В частности для независимых опытов:

$$p_{a_1 a_2 \dots a_m}(a_{m+1}) = p(a_{m+1}),$$

для простой цепи:

$$p_{a_1 a_2 \dots a_m}(a_{m+1}) = p_{a_m}(a_{m+1}),$$

для двусложной цепи:

$$p_{a_1 a_2 \dots a_m}(a_{m+1}) = p_{a_{m-1}, a_m}(a_{m+1})$$

и т. д.

Введем в рассмотрение величины:

$$H_0 = - \sum_{a_1=1}^n p(a_1) \cdot \lg p(a_1)$$

$$H_1 = - \sum_{a_1=1}^n \sum_{a_2=1}^n p(a_1) p_{a_1}(a_2) \cdot \lg p_{a_1}(a_2)$$

$$H_2 = - \sum_{a_1=1}^n \sum_{a_2=1}^n \sum_{a_3=1}^n p(a_1) \cdot p_{a_1}(a_2) \cdot p_{a_1 a_2}(a_3) \lg p_{a_1 a_2}(a_3)$$

и вообще

$$H_\mu = - \sum_{a_1} \sum_{a_2} \dots \sum_{a_\mu} p(a_1) p_{a_1}(a_2) \dots$$

$$\dots p_{a_1 a_2 \dots a_\mu}(a_{\mu+1}) \lg p_{a_1 \dots a_\mu}(a_{\mu+1}). \quad (1.1)$$

Величина H_μ называется одношаговой энтропией цепи Маркова; очевидно, что для независимых опытов

$$H_\mu = H_{\mu-1} = \dots = H_0,$$

для простой цепи

$$H_\mu = H_{\mu-1} = \dots = H_1,$$

для двусложной цепи

и т. д.

$$H_\mu = H_{\mu-1} = \dots = H_2$$

2. Объединим r опытов и рассмотрим соответствующую конечную схему с состояниями:

$$A_{a_1} A_{a_2} \dots A_{a_r}$$

и соответствующими вероятностями:

$$p(a_1, a_2, \dots, a_r)$$

Величину энтропии этой схемы, осередненную по числу опытов

$$H_r = - \frac{1}{r} \sum_{a_1} \sum_{a_2} \dots \sum_{a_r} p(a_1, a_2, \dots, a_r) \lg p(a_1, \dots, a_r) \quad (1.2)$$

будем называть осередненной одношаговой энтропией.

Для цепи Маркова порядка μ при $r > \mu$ будем иметь:

или окончательно

$$\bar{H}_r = \frac{1}{r} [H_0 + H_1 + H_2 + \dots + H_{r-1} + (r - \mu) H_\mu] \quad (1.3)$$

При $r \rightarrow \infty$, имеем

$$\lim_{r \rightarrow \infty} \bar{H}_r = H_\mu. \quad (1.4)$$

Таким образом одношаговую энтропию цепи Маркова, порядка μ можно рассматривать, как предел осередненной одношаговой энтропии, когда число опытов стремится к бесконечности.

3. Допустим, что в цепи Маркова порядка μ при первых μ опытах имели место события

$$A_{x_1}, A_{x_2}, \dots, A_{x_\mu};$$

составим энтропию конечной схемы, представляющей объединение следующих за этими опытами, r опытов, т. е. величину:

$$H_{x_1 x_2 \dots x_\mu}^{(r)} = - \sum_{x_{\mu+1}} \sum_{x_{\mu+2}} \dots \dots \sum_{x_{\mu+r}} p_{x_1 x_2 \dots x_r} (x_{\mu+1}, \dots, x_{\mu+r}) \lg p_{x_1 x_2 \dots x_r} (x_{\mu+1}, \dots, x_{\mu+r});$$

тогда величина

$$H^{(r)} = \sum_{x_1} p(x_1) \sum_{x_2} p_{x_1}(x_2) \dots \dots \sum_{x_\mu} p_{x_1 x_2 \dots x_{\mu-1}}(x_\mu) H_{x_1 x_2 \dots x_\mu}^{(r)} \quad (1.5)$$

называется r -шаговой энтропией цепи Маркова порядка μ ; очевидно $H^{(1)} = H_\mu$.

Осередненная одношаговая энтропия $\bar{H}_r (r > \mu)$ выразится через многошаговую энтропию следующим образом:

$$\begin{aligned} \bar{H}_r &= - \frac{1}{r} \sum_{x_1} \sum_{x_2} \dots \sum_{x_r} p(x_1, x_2, \dots, x_r) \lg p(x_1, x_2, \dots, x_r) = \\ &= - \frac{1}{r} \sum_{x_1} p(x_1) \lg p(x_1) \sum_{x_2} \dots \sum_{x_r} p_{x_1}(x_2, \dots, x_r) - \\ &- \frac{1}{r} \sum_{x_1} p(x_1) \sum_{x_2} p_{x_1}(x_2) \lg p_{x_1}(x_2) \sum_{x_3} \dots \sum_{x_r} p_{x_1 x_2}(x_3, \dots, x_r) - \end{aligned}$$

* Величина $H^{(r)}$ зависит от μ , но т. к. мы здесь будем рассматривать ее при одном и том же μ , то эту зависимость от μ отмечать не будем, чтобы не усложнять обозначений.

$$\begin{aligned}
& -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \sum_{\alpha_3} p_{\alpha_1 \alpha_2}(\alpha_3) \lg p_{\alpha_1 \alpha_2}(\alpha_3) \sum_{\alpha_4} \cdots \sum_{\alpha_r} p_{\alpha_1 \alpha_2 \alpha_3}(\alpha_4, \dots, \alpha_r) - \\
& \dots - \\
& -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \cdots \sum_{\alpha_\mu} p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_\mu) \lg p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_\mu) \sum_{\alpha_{\mu+1}} \cdots \\
& \cdots \sum_{\alpha_r} p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}, \dots, \alpha_r) \\
& -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \cdots \sum_{\alpha_\mu} p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_\mu) \sum_{\alpha_{\mu+1}} \cdots \\
& \cdots \sum_{\alpha_r} p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}, \dots, \alpha_r) \lg p_{\alpha_1 \dots \alpha_\mu}(\alpha_{\mu+1}, \dots, \alpha_r) = \\
& = -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \lg p(\alpha_1) - \frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p(\alpha_2) \lg p_{\alpha_1}(\alpha_2) - \\
& -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \sum_{\alpha_3} p_{\alpha_1 \alpha_2}(\alpha_3) \lg p_{\alpha_1 \alpha_2}(\alpha_3) - \\
& \dots - \\
& -\frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \sum_{\alpha_3} \cdots \\
& \cdots \sum_{\alpha_\mu} p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_\mu) \lg p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_\mu) - \\
& + \frac{1}{r} \sum_{\alpha_1} p(\alpha_1) \sum_{\alpha_2} p_{\alpha_1}(\alpha_2) \sum_{\alpha_3} \cdots \\
& \cdots \sum_{\alpha_\mu} p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}} H_{\alpha_1 \alpha_2 \dots \alpha_\mu}^{(r-\mu)}
\end{aligned}$$

или

$$\bar{H}_r = \frac{1}{r} (H_0 + H_1 + H_2 + \cdots + H_{\mu-1} + H^{(r-\mu)})$$

или окончательно заменяя r через $r + \mu$, получаем:

$$\bar{H}_{r+\mu} = \frac{1}{r+\mu} (H_0 + H_1 + H_2 + \cdots + H_{\mu-1} + H^{(r)}). \quad (1.6)$$

Согласно формуле (1.3), имеем с другой стороны:

$$\bar{H}_{r+\mu} = \frac{1}{r+\mu} (H_0 + H_1 + H_2 + \dots + H_{\mu-1}) + \frac{r}{r+\mu} H_{\mu}$$

следовательно

$$H^{(r)} = rH_{\mu} \quad (1.7)$$

Этот результат, другим методом при $\mu = 1$, был получен в работе [1].

II. Предельная теорема для энтропии сложной цепи Маркова порядка μ

Предельная теорема Хинчина [1] для энтропии простой цепи Маркова, легко обобщается на сложные цепи. Допустим, что сложная цепь Маркова порядка μ подчиняется закону больших чисел, т. е. при произвольных $\delta > 0$, $\varepsilon > 0$ и достаточно большом s :

$$P \left\{ \left| \frac{m_i}{s} - p(i) \right| > \delta \right\} < \varepsilon, \quad (i = 1, 2, \dots, n)$$

где $m_1, m_2, \dots, m_n$ числа попаданий соответственно в состояния $A_1, A_2, \dots, A_n$. Всякий возможный результат s опытов может быть записан в виде «цепочки C »

$$C = A_{k_1} A_{k_2} \dots A_{k_s}$$

где числа $k_1, k_2, \dots, k_s$ числа из $1, 2, \dots, n$. Вероятность осуществления цепочки C будет:

$$P(C) = p(k_1) p_{k_1}(k_2) p_{k_1 k_2}(k_3) \dots \\ \dots p_{k_1 k_2 \dots k_{\mu}}(k_{\mu+1}) p_{k_1 k_2 \dots k_{\mu+1}}(k_{\mu+2}) \dots p_{k_1 k_2 \dots k_{s-\mu}}(k_{s-\mu+1}) \dots p_{k_1 k_2 \dots k_{s-1}}(k_s),$$

которую можно записать в виде:

$$P(C) = p(k_1) p_{k_1}(k_2) \dots p_{k_1 k_2 \dots k_{\mu-1}}(k_{\mu}) \prod_{\alpha_1, \alpha_2} \dots \\ \dots \prod_{\alpha_{\mu+1}} [p_{\alpha_1 \alpha_2 \dots \alpha_{\mu}}(\alpha_{\mu+1})]^{m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}}, \quad (2.1)$$

где числа $m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}$ представляют число случаев, в написанном произведении для $P(C)$, при котором:

$$k_m = \alpha_1, k_{m+1} = \alpha_2, \dots, k_{m+\mu} = \alpha_{\mu+1}.$$

Теорема: при произвольных $\varepsilon > 0$, $\eta > 0$ и достаточно большом s , все цепочки C могут быть разбиты на две группы, обладающие свойствами:

1) вероятность $P(C)$ любой цепочки первой группы удовлетворяет неравенству:

$$\left| \frac{\lg \frac{1}{P(C)}}{s} - H_{\mu} \right| < \eta$$

или

$$a^{-s(H+\eta)} < P(C) < a^{-s(H-\eta)}$$

2) сумма вероятностей всех цепочек второй группы меньше чем ε .

Доказательство. Разобьем все цепочки S длины s на две группы: цепочка принадлежит первой группе, если она возможна, т. е. $P(C) > 0$ и если для любых $\alpha_1, \alpha_2, \dots, \alpha_\mu, \alpha_{\mu+1}$ ($1 \leq \alpha_i \leq n$) выполняются неравенства:

$$|m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} - sp(\alpha_1)p_{\alpha_1}(\alpha_2) \dots p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1})| < s\delta; \quad (2.2)$$

ко второй группе отнесем все остальные цепочки. Докажем, что это разбиение при достаточно большом s удовлетворяет условиям теоремы. Допустим сначала, что цепочка s первой группы; тогда

$$\begin{aligned} m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} &= sp(\alpha_1)p_{\alpha_1}(\alpha_2)p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) + \\ &+ s\delta\theta_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} \\ |\theta_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}| &< 1. \end{aligned}$$

Эти выражения для $m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}$ внесем в (2.1), причем требование возможности данной цепочки, необходимо, влечет за собой $m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} = 0$ при $p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) = 0$, вследствие чего в произведении (2.1) мы должны ограничиться множителями для которых

$$p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) > 0,$$

что будем отмечать * при произведениях. Имеем:

$$P(C) = p(k_1)p_{k_1}(k_2) \dots p_{k_1 k_2 \dots k_{\mu-1}}(k_\mu).$$

$$\prod_{\alpha_1} \prod_{\alpha_2} \dots \prod_{\alpha_{\mu+1}}^* [p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1})]^{sp(\alpha_1)p_{\alpha_1}(\alpha_2) \dots p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) + s\delta\theta_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}}}$$

или

$$\begin{aligned} \lg \frac{1}{P(C)} &= -\lg p(k_1) - \lg p_{k_1}(k_2) - \dots - \lg p_{k_1 k_2 \dots k_{\mu-1}}(k_\mu) - \\ &- s \sum_{\alpha_1} \sum_{\alpha_2} \dots \sum_{\alpha_{\mu+1}}^* p(\alpha_1)p_{\alpha_1}(\alpha_2) \dots p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) \lg p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}) - \\ &- s\delta \sum_{\alpha_1} \sum_{\alpha_2} \dots \sum_{\alpha_{\mu+1}}^* \theta_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} \lg p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1}); \end{aligned}$$

отсюда

$$\begin{aligned} \lg \frac{1}{P(C)} - H_s &= \frac{1}{s} \left(\lg \frac{1}{p(k_1)} + \lg \frac{1}{p_{k_1}(k_2)} + \dots + \lg \frac{1}{p_{k_1 k_2 \dots k_{\mu-1}}(k_\mu)} \right) + \\ &+ \delta \sum_{\alpha_1} \sum_{\alpha_2} \dots \sum_{\alpha_{\mu+1}}^* \theta_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} \lg \frac{1}{p_{\alpha_1 \alpha_2 \dots \alpha_\mu}(\alpha_{\mu+1})} \end{aligned}$$

и

$$\left| \frac{\lg \frac{1}{P(C)}}{s} - H_{\mu} \right| < \frac{1}{s} \left(\lg \frac{1}{p(k_1)} + \lg \frac{1}{p_{k_1}(k_2)} + \dots \right. \\ \left. \dots + \lg \frac{1}{p_{k_1 k_2 \dots k_{\mu-1}}(k_{\mu})} \right) + \delta \sum_{\alpha_1} \sum_{\alpha_2} \dots \sum_{\alpha_{\mu+1}} \lg \frac{1}{p_{\alpha_1 \alpha_2 \dots \alpha_{\mu}}(\alpha_{\mu+1})}$$

и следовательно при достаточно больших s :

$$\left| \frac{\lg \frac{1}{P(C)}}{s} - H_{\mu} \right| < \eta.$$

Рассмотрим теперь сумму вероятностей всех цепочек второй группы; заметим, что невозможные цепочки могут не учитываться, т. к. вероятности их равны нулю; здесь нужно оценить сумму вероятностей тех цепочек, для которых хотя бы при одном наборе индексов $\alpha_1, \alpha_2, \dots, \alpha_{\mu+1}$ нарушается неравенство (2.2), а она очевидно меньше чем

$$\sum_{\alpha_1} \sum_{\alpha_2} \dots \sum_{\alpha_{\mu+1}} P \{ |m_{\alpha_1 \alpha_2 \dots \alpha_{\mu+1}} - sp(\alpha_1) p_{\alpha_1}(\alpha_2) \dots p_{\alpha_1 \alpha_2 \dots \alpha_{\mu}}(\alpha_{\mu+1})| \geq s\delta \}$$

Зафиксируем индексы $\alpha_1, \alpha_2, \dots, \alpha_{\mu+1}$; в силу закона больших чисел — при достаточно малых $\varepsilon > 0$, $\delta > 0$ и достаточно большом s :

$$P \left\{ \left| m_{\alpha_1} - sp(\alpha_1) \right| < \frac{\delta}{\mu+1} s \right\} > 1 - \varepsilon; \quad (2.3)$$

если неравенство в скобках выполнено, то m_{α_1} достаточно велико и поэтому в силу закона больших чисел:

$$P \left\{ \left| \frac{m_{\alpha_1 \alpha_2}}{m_{\alpha_1}} - p_{\alpha_1}(\alpha_2) \right| < \frac{\delta}{\mu+1} \right\} > 1 - \varepsilon$$

или

$$P \left\{ \left| m_{\alpha_1 \alpha_2} - m_{\alpha_1} p_{\alpha_1}(\alpha_2) \right| < \frac{\delta}{\mu+1} m_{\alpha_1} < \frac{\delta}{\mu+1} s \right\} > 1 - \varepsilon; \quad (2.4)$$

при выполнении неравенства в скобках, имеем точно также

$$P \left\{ \left| m_{\alpha_1 \alpha_2 \alpha_3} - m_{\alpha_1 \alpha_2} p_{\alpha_1 \alpha_2}(\alpha_3) \right| < \frac{\delta}{\mu+1} s \right\} > 1 - \varepsilon; \quad (2.5)$$

аналогично рассуждая получим еще, ряд неравенств:

$$P \left\{ \left| m_{\alpha_1 \alpha_2 \alpha_3 \alpha_4} - m_{\alpha_1 \alpha_2 \alpha_3} p_{\alpha_1 \alpha_2 \alpha_3}(\alpha_4) \right| < \frac{\delta}{\mu+1} s \right\} > 1 - \varepsilon \quad (2.5)$$

$$\dots \dots \dots \quad (2.6)$$

$$P \left\{ \left| m_{\alpha_1 \alpha_2 \dots \alpha_{\mu}} - p_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}}(\alpha_{\mu}) m_{\alpha_1 \alpha_2 \dots \alpha_{\mu-1}} \right| < \frac{\delta}{\mu+1} s \right\} > 1 - \varepsilon$$

ношаговая энтропия, а $H^{(s)}$, s —шаговая энтропия, соответствующие рассматриваемой цепи Маркова порядка μ .

Ереванский политехнический институт
им. К. Маркса

Поступило 21 I 1958

Փ. Հ. Համբարձումյան

ՄԱՐԿՈՎԻ ՇՂԹԱՆԵՐԻ ԷՆՏՐՈՊԻԱՅԻ ՄԱՍԻՆ

Ա Մ Փ Ո Փ Ո Ւ Մ

Աշխատության առաջին մասում ընդհանրացվում են էնտրոպիայի տարրեր տեսակի գաղափարները՝ Մարկովի բարդ շղթաների համար, և ստացվում են այդ տարրեր գաղափարների կապերը միմյանց հետ: Երկրորդ մասում էնտրոպիայի վերաբերյալ սահմանալիս թևորումը Մարկովի պարզ շղթաների համար ընդհանրացված է՝ Մարկովի ցանկացած կարգի բարդ շղթաների համար:

Л И Т Е Р А Т У Р А

1. Хинчин А. Я. Понятие энтропии в теории вероятностей. Усп. мат. наук, том, VIII, вып. 3, (1953).