ОЦЕНКИ КРИВЫХ НАУЧЕНИЯ

М.В. Белов, Д.А. Новиков, А.Д. Рогаткин

Получены верхние и нижние оценки кривых научения для моделей индивидуального и совместного научения в дискретном и непрерывном времени.

Ключевые слова: модель научения, кривая научения, совместное научение.

1. Введение

В книге [1] предложена достаточно общая (включающая многие известные) вероятностная модель научения в дискретном – см. выражение (1) ниже - и непрерывном – см. выражение (19) - времени. Целью настоящей работы является получение простых предельных/асимптотических (по значениям параметров модели или времени) верхних и нижних оценок (в основном, экспоненциальных) кривых научения.

Во втором разделе в соответствии с [1] кратко описана модель научения в дискретном времени, оценки соответствующих кривых научения приведены в третьем разделе. В четвертом разделе сравниваются одноуровневая и двухуровневая схемы научения.

В пятом разделе рассмотрена модель научения в непрерывном времени, которая в шестом разделе обобщена на случай совместного научения нескольких субъектов, для которого получены оценки соответствующих кривых научения. Седьмой раздел посвящен агрегированному описанию совместного научения, восьмой – сетевым моделям научения.

Наконец, девятый раздел содержит стохастические оценки значений уровня научения.

2. Дискретная модель научения

Пусть *обучаемый* должен освоить новый для него вид *деятельности*, имеющий конечное число $K \geq 2$ независимых *компонентов*. Процесс научения заключается в предъявлении обучаемому последовательно (в моменты времени $t = 1, 2, \ldots$), но случайным образом, *ситуаций*, в каждой из которых, предъявленных впервые, он должен освоить один определенный компонент деятельности (между ситуациями и компонентами деятельности существует взаимно однозначное соответствие). Введем следующие предположения (описываемую ими модель называют *базовой моделью научения*).

I. До начала научения обучаемый не встречался ни с одной из ситуаций. Ситуации предъявляются ему по одной случайным образом, причем в любой момент времени вероятность предъявления конкретной ситуации постоянна и не зависит от предыстории (в терминах теории вероятностей: рассматривается K-цветная урновая схема с возвращением).

II. Каждый компонент деятельности может быть либо уже освоен, либо пока не освоен.

III. Если предъявляется ситуация, еще не встречавшаяся обучаемому ранее в процессе научения, то он на текущем шаге гарантированно осваивает соответствующий этой ситуации компонент своей деятельности.

IV. Если предъявляется ситуация, уже встречавшаяся обучаемому ранее в процессе научения, то он успешно идентифицирует соответствующую этой ситуации компонент своей деятельности (забывания не происходит).

V. Значением *критерия уровня научения* является вероятность того, что обучаемому будет предъявлена уже встречавшаяся ранее ситуация (математическое ожидание доли освоенных ситуаций).

Таким образом, в соответствии с предположениями I-V:

- *начальное значение* (в нулевой момент времени) уровня научения равно нулю;
- *кривая научения* (последовательность значений уровня научения) не убывает и асимптотически стремится к единице.

Обозначим через $p_k > 0$ вероятность того, что на очередном шаге обучаемому будет предъявлена k-я ситуация (очевидно, $\sum_{k=1}^{K} p_k = 1$). Вектор этих вероятностей обозначим через $P = (p_1, \ldots, p_K)$, значение критерия *уровня научения* в момент времени t – через x_t.

В [1] доказано, что

(1) $x_t(P) = 1 - \sum_{k=1}^{K} p_k (1 - p_k)^t$, $t = 0, 1, 2, \ldots$.

Для кривой научения (1) имеет место следующие достаточно грубые экспоненциальные оценки:

(2) $1 - e^{\gamma^- t} \le x_t(P) \le 1 - e^{\gamma^+ t}$,

где $\gamma^- = -\ln(1 - \min_k \{p_k\})$, $\gamma^+ = -\ln(1 - \max_k \{p_k\})$.

Введем параметр $\rho \in [0; 1/K]$. Обозначим через

(3) $P_{\rho, K} = \{P = (p_1, \ldots, p_K) \mid \sum_{k=1}^{K} p_k = 1, p_k \ge \rho, k = \overline{1,K}\}$

множество K-мерных распределений вероятностей, значение каждой из которых не менее *порога* ρ.

В [1] доказано, что максимум выражения (1) по всевозможным распределениям вероятностей $P \in P_{\rho, K}$ достигается на равномерном распределении. Подставляя $p_k = 1/K$, $k = \overline{1,K}$ в выражение (1), получим

(4) $x_t^*(K) = 1 - \left(1 - \frac{1}{K}\right)^t = 1 - \exp(-\gamma t)$,

где $\gamma(K) = \ln(1 + 1/(K - 1))$ - *скорость научения*.

3. Оценки «дискретных» кривых научения

Рассмотрим оценки кривых научения в зависимости от распределений $P \in P_{\rho, K}$ и значений параметров ρ и K. Начнем с оценки снизу.

Утверждение 1. $\forall\ K \ge 2$, $\forall \rho \in (0; 1/K]$ $\exists t^*(\rho) = \frac{2}{\rho} - 1$ такое, что $\forall \tau > t^*(\rho)$ решение $p^{\min}$ задачи

(5) $x_\tau(P) \to \min_{P \in P_{\rho, K}}$

имеет вид:

(6) $p_1^{\min} = 1 - (K - 1)\rho$, $p_k^{\min} = \rho$, $k = \overline{2,K}$.

Доказательство. В [1] доказана

Лемма 1. $\forall \rho \in (0; 1/K]$ $\exists t^*(\rho)$ такое, что $\forall \tau > t^*(\rho)$ $\sum_{k=1}^{K} p_k (1 - p_k)^\tau$ – строго выпуклая функция $\{p_k\}_{k=\overline{1,K}}$ (оценка $t^*(\rho) = \frac{2}{\rho} - 1$ получается непосредственно из условия положительности второй производной слагаемых).

Строго выпуклая функция достигает на ограниченном выпуклом множестве своего максимума в одной из крайних точек этого множества. Точка $p^{\min}$ является крайней точкой выпуклого многогранника $P_{\rho, K}$. В силу симметрии целевой функции, минимум последней достигается в том числе и в этой точке (отметим, что значения во всех K крайних точках одинаковы). Утверждение 1 доказано.

Воспользовавшись результатом (6) утверждения 1, вычислим при $\tau > t^*(\rho)$ оценку снизу величины (1):

(7) $x_\tau^{\min}(\rho,K)=1-(K-1)^\tau\rho^\tau[1-(K-1)\rho]-(K-1)\rho(1-\rho)^\tau$.

Найдем теперь оценку кривой научения сверху. В [1] показано, что при $\tau>t^*(\rho)$ максимум $x_\tau(P)$ достигается при равномерном распределении вероятностей реализации ситуаций. Подставляя в (1) $p_k=\frac{1}{K}$, $k=\overline{1,K}$, вычислим соответствующую оценку сверху: $x_\tau^{\max}(\rho,K)=1-(1-\frac{1}{K})^\tau$ (отметим, что при $0\le\tau\ll t^*(\rho)$ имеет место $x_\tau^{\min}(\rho,K)\ge x_\tau^{\max}(\rho,K)$ - см. Рис. 1). Преобразовывая, получим

(8) $x_\tau^{\max}(\rho,K)=1-\exp\left(-\gamma(K)\tau\right)$.

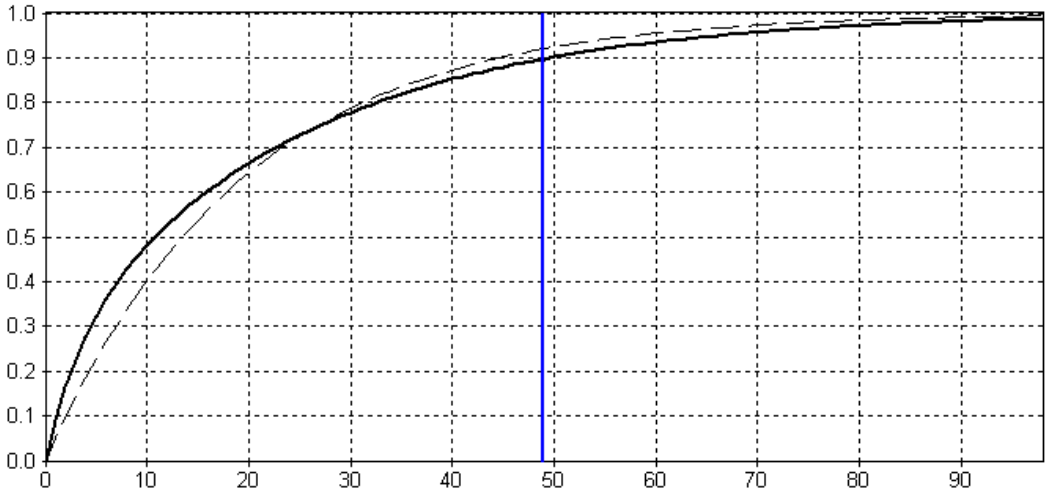

Рис. 1. Зависимости $x_\tau^{\min}(\rho,K)$ (сплошная линия) и $x_\tau^{\max}(\rho,K)$ (пунктирная линия) при K = 20, ρ = 0,04. Вертикальная линия - $t^*(\rho)=49$

Так как $\forall\ K\ge 2\ \ \forall\rho\in(0;1/K]$, $\forall\ P\in P_{\rho,K}$, $\forall\tau>t^*(\rho)\ \ x_\tau^{\min}(\rho,K)\le x_\tau(P)\le x_\tau^{\max}(\rho,K)$ и $\lim\limits_{\tau\to+\infty}x_\tau^{\min}(\rho,K)=x_\tau^{\max}(\rho,K)=1$, то $\lim\limits_{\tau\to+\infty}x_\tau(P)=1$ и $\lim\limits_{\tau\to+\infty}(x_\tau^{\max}(\rho,K)-x_\tau^{\min}(\rho,K))=0$. Итак, мы обосновали справедливость следующего утверждения.

Утверждение 2. $\forall\ K\ge 2$, $\forall\rho\in(0;1/K]$, $\forall\ P\in P_{\rho,K}$, $\forall\tau>t^*(\rho)$ выполнено:

1) $x_\tau^{\min}(\rho,K)\le x_\tau(P)\le x_\tau^{\max}(\rho,K)$;
2) $x_\tau^{\min}(\rho,K)$ возрастает по $\rho\in(0;1/K]$;
3) $x_\tau^{\min}(\rho,K)$ и $x_\tau^{\max}(\rho,K)$ убывают по K;
4) $x_\tau^{\min}(\rho,K)\le x_\tau^{\max}(\rho,K)$;
5) $x_\tau^{\min}(\frac{1}{K},K)=x_\tau^{\max}(\frac{1}{K},K)$;
6) $\lim\limits_{\tau\to+\infty}(x_\tau^{\max}(\rho,K)-x_\tau^{\min}(\rho,K))=0$.

Обозначим $\beta(\rho)=\ln(\frac{1}{1-\rho})$. Рассмотрим выражение (7). Так как при $\rho\le\frac{1}{K}$ имеет место $(K-1)\rho\le 1$, то $x_\tau^{\min}(\rho,K)\ge(K-1)\rho(1-(1-\rho)^\tau)=(K-1)\rho(1-e^{-\beta(\rho)\tau})$. Помимо точных оценок (7) и (8), получаем следующую удобную для практики приблизительную оценку снизу $x_\tau(P)$ при $\tau>t^*(\rho)$: $\forall\ P\in P_{\rho,K}$

$$1-(1-\rho)^\tau\le x_\tau(P)\le 1-(1-\frac{1}{K})^\tau,$$

или

$$1-e^{-\beta(\rho)\tau} \le x_\tau(P) \le 1-e^{-\gamma(K)\tau}.$$

Фиксируем произвольное распределение $P \in P_{\rho,K}$. Обозначим через P^α его линейную комбинацию с равномерным распределением: $p_k^\alpha = \frac{\alpha}{K} + (1-\alpha)p_k$, $k=\overline{1,K}$, $\alpha \in [0;1]$. Из леммы 1 следует справедливость следующего утверждения.

Утверждение 3а. $\forall \rho \in (0;1/K]$ $\exists t^*(\rho)$ такое, что $\forall \tau > t^*(\rho)$ $\forall P \in P_{\rho,K}$ $x_\tau(P^\alpha) \ge x_\tau(P)$ и $x_\tau(P^\alpha)$ не убывает по $\alpha \in [0;1]$.

Эффективность научения и большое число агентов. Отказавшись от предположения III, в настоящем подразделе будем считать, что при предъявлении *k*-ой ситуации успешное освоение соответствующего компонента деятельности происходит с постоянной вероятностью $w_k \le 1$, (не зависящей от времени, достигнутого уровня научения и т.д.). Вектор этих вероятностей – который можно условно назвать *эффективностью научения* - обозначим через $W = (w_1, \ldots, w_K)$. Отметим, что вектор W в общем случае не удовлетворяет условию нормировки.

Тогда кривая научения примет вид

(9) $x_t(P,W) = \sum_{k=1}^{K} p_k \left(1-\left(1-w_k p_k\right)^t\right)$, $t = 0, 1, 2, \ldots$.

Отметим, что выражение (1) является частным случаем выражения (9), соответствующим единичным вероятностям научения при предъявлении любой ситуации: $w_k = 1$, $k=\overline{1,K}$. В качестве частного случая (9) можно привести ситуацию, когда осваивается единственный компонент деятельности ($K = 1$, $p_1 = 1$), тогда из (9) следует, что $x_t(w_1) = 1-\left(1-w_1\right)^t = 1-\exp\left(-\gamma_w t\right)$, где $\gamma_w = -\ln(1-w_1)$.

В частном случае равномерного распределения выражение (9) примет вид:

(10) $x_t(P,W) = 1-\frac{1}{K}\sum_{k=1}^{K}\left(1-\frac{w_k}{K}\right)^t$.

Фиксируем произвольный вектор W. Обозначим через W^α его линейную комбинацию с единичным вектором: $w_k^\alpha = \alpha + (1-\alpha)w_k$, $k=\overline{1,K}$, $\alpha \in [0;1]$. Из леммы 1 следует справедливость следующего утверждения.

Утверждение 3б. $\forall \rho \in (0;1/K]$ $\exists t^*(\rho)$ такое, что $\forall \tau > t^*(\rho)$ $\forall W$ $x_\tau(P,W^\alpha) \ge x_\tau(P,W)$ и $x_\tau(P,W^\alpha)$ не убывает по $\alpha \in [0;1]$.

Содержательным следствием утверждения 3б является то, что при нормированных эффективностях научения максимум уровня научения (10) будет достигаться при равных вероятностях успешного освоения всех компонентов деятельности.

4. Иерархическая декомпозиция компонентов деятельности

Сравним две схемы научения: *одноуровневую*, в которой происходит освоение K компонент деятельности, и *двухуровневую*, в которой сначала на верхнем уровне субъект осваивает (учится различать) на верхнем уровне $m \le K$ ситуаций, а на затем нижнем уровне – параллельно для каждой из m компонент верхнего уровня – осваивает ее «детализацию» на K/m «подкомпонент» (понятно, что, если подкомпоненты, соответствующие различным компонентам верхнего уровня, осваиваются последовательно, то двухуровневая схема менее эффективна). В обеих схемах число возможных компонент нижнего уровня одно и то же – K.

Двухуровневая схема научения будет предпочтительнее одноуровневой, если, во-первых, уровень научения в ней (определяемый как вероятность того, что на обоих уровнях обучаемому встретятся уже встречавшиеся ранее ситуации) выше, т.е. должно выполняться

(11) $x_t^*(m)\ \ x_t^*(K/m) \geq x_t^*(K)$.

Подставляя выражение (4) в выражение (11), легко убедиться, что максимум левой части (11) достигается при $m^*(K) = \sqrt{K}$, т.е. оптимальна симметричная декомпозиция. При этом неравенство (11), принимающее вид

$$1-\left(1-\frac{1}{\sqrt{K}}\right)^t \geq \sqrt{1-\left(1-\frac{1}{K}\right)^t},$$

справедливо.

Во-вторых, необходимо проверить, что последовательное освоение компонентов верхнего и нижнего уровня требует не больше времени, чем в одноуровневой схеме научения. В [1] показано, что минимальное время, требуемое для достижения заданного уровня научения $x \in [0;1]$ при числе компонентов деятельности K равно $t(x, K) = \frac{\ln(1-x)}{\ln(1-\frac{1}{K})}$. Для того, чтобы достичь общего уровня научения x, на каждом из двух уровней должен быть достигнут уровень научения $\sqrt{x}$. Получаем следующее требование на соотношение времен:

(12) $t(\sqrt{x}, K/m) + t(x\sqrt{x}, K/m) \leq t(x, K)$.

Подставляя в (12) $m = m^*(K)$, получим

(13) $\frac{2\ln(1-\sqrt{x})}{\ln(1-\frac{1}{\sqrt{K}})} \leq \frac{\ln(1-x)}{\ln(1-\frac{1}{K})}$.

Разлагая в ряд Тейлора, получаем из выражения (13) следующее требование на соотношение параметров: $x \geq 4/K$. При $K < 4$ декомпозиция не имеет смысла. Самая простая декомпозиция при $K = 4$ (две ситуации на вернем уровне, каждой из которых соответствуют по две подситуации на нижнем) требует достижения единичного уровня научения на каждом из уровней, что нереалистично в рамках рассматриваемой вероятностной модели. А вот при бо́льших значениях требования на уровни научения вполне разумны: например, при $K = 100$ оптимальным является разбиение множества возможных ситуаций на 10 групп по 10 ситуаций в каждой, с достижением уровня научения 0,2 на верхнем и нижнем уровне.

Подставляя условие $x \geq 4/K$ в левую часть выражения (13), получим оценку минимальной продолжительности научения, при котором целесообразна иерархическая декомпозиция множества возможных ситуаций:

(14) $t_{\min}(K) = \frac{2\ln(1-\frac{2}{\sqrt{K}})}{\ln(1-\frac{1}{\sqrt{K}})}$.

Отметим, что до сих пор мы сравнивали одноуровневую и двухуровневую схемы научения. Условия эффективности последней могут быть применены и к ней самой, и т.д. Поэтому перспективным, но достаточно простым с «технической» точки зрения, представляется получение оценок для оптимального числа уровней симметричной (а, судя по всему, именно она оптимальна) иерархической декомпозиции множества возможных ситуаций.

5. Модель научения в непрерывном времени

Рассмотрим, следуя [1], общую *модель научения* в *непрерывном времени* - дифференциальное уравнение для уровня научения $x(t) \in [0;1], t \geq 0$:

(15) $\dot{x}(t) = (1-x)\, f(x)$

с начальным значением $x(0) \in [0; 1)$, где $f(\cdot)$: $[0; 1] \to (0; A]$ – непрерывная функция, где $0 < A < +\infty$ (если значение f интерпретируется как вероятность или как значение уровня научения в другом процессе научения – см. [1], то необходимо потребовать чтобы $A = 1$).

В [1] показано, что из вида правой части уравнения (15) и введенных предположений следует, что:

а) решение уравнения (15) существует и единственно;

б) кривая научения $x(t)$ является строго монотонно возрастающей и $\forall t \geq 0 \;\; \dot{x}(t) \leq A$, т.е. скорость ее роста ограничена;

в) кривая научения мажорируется экспоненциальной кривой:

(16) $\forall t \geq 0 \;\; x(t) \leq 1 - (1 - x(0)) \exp(-A\,t)$,

г) кривая научения $x(t)$ является *замедленно-асимптотической*, т.е.

$$\lim_{t \to +\infty} x(t) = 1, \;\; \lim_{t \to +\infty} \dot{x}(t) = 0 .$$

Варьируя функцию $f(\cdot)$, можно получать различные кривые научения. Частными случаями решения уравнения (15) являются экспоненциальные, степенные, логистические и другие хрестоматийные классы кривых научения [1].

6. Совместное научение

Пусть имеется $n \geq 1$ «обучаемых» - *агентов*, которые могут представлять собой либо отдельных субъектов, либо метакомпоненты деятельности одного субъекта. Обозначая через i номер агента, через $x_i \in [0;1]$ – уровень его научения, через $x_i(0) \in [0;1)$ – начальное значение уровня его научения, через $X = (x_1, x_2, \ldots, x_n)$ – вектор уровней научения, запишем для каждого из агентов аналог уравнения (15):

(17) $\dot{x}_i(t) = (1 - x_i)\, f_i(X), \;\; i = \overline{1,n}$.

Относительно *функций взаимного влияния агентов* $f_i(\cdot)$: $[0; 1]^n \to (0; A_i]$ предположим, что:

А.1. они непрерывно дифференцируемы по всем переменным, и $0 < A_i < +\infty$ (если значение f_i интерпретируется как вероятность или как значение уровня научения в другом процессе научения – см. [1], то необходимо потребовать чтобы $A_i = 1$);

А.2. $\forall X \in [0;1]^n \;\; \forall i, j = \overline{1,n}$ частные производные $\dfrac{\partial f_i(X)}{\partial x_j}$ знакопостоянны.

Отсюда следует, что, во-первых, кривые научения агентов удовлетворяют вышеприведенным свойствам а)-г). Во-вторых, последнее предположение позволяет получить более «тонкие» оценки, нежели выражения типа (16). Обозначим

$$\xi_{ij} = \max\{\mathrm{Sign}(\frac{\partial f_i(X)}{\partial x_j}); x_j(0)\}, \;\; \zeta_{ij} = \max\{-\mathrm{Sign}(\frac{\partial f_i(X)}{\partial x_j}); x_j(0)\},$$

$$\mu_i^- = (\zeta_{i1}, \ldots, \zeta_{in}), \;\; \mu_i^+ = (\xi_{i1}, \ldots, \xi_{in}), \;\; i, j = \overline{1,n}.$$

Утверждение 4. В рамках предположений А.1-А.2 справедливы следующие оценки кривых научения $\{x_i(t)\}$, являющихся решениями системы дифференциальных уравнений (17):

$$\forall i = \overline{1,n} \;\; \forall t \geq 0 \;\; x_i^{\min}(t) \leq x_i(t) \leq x_i^{\max}(t),$$

где

(18) $x_i^{\min}(t) = 1 - (1 - x_i(0)) \exp(-f_i(\mu_i^-)\, t)$,

(19) $x_i^{\max}(t) = 1 - (1 - x_i(0)) \exp(-f_i(\mu_i^+)\, t)$.

Справедливость утверждения 4 следует из того, что в рамках предположения А.2 непрерывные монотонные (по соответствующим переменным) функции $\{f_i(\cdot)\}$ достигают своих минимумов и максимумов в вершинах куба $\prod_{i=1}^{n}[x_i(0);1]$.

7. Агрегированное описание

Рассмотрим частный случай системы дифференциальных уравнений (17), когда $f_i(X) = \gamma_i G(X)$, $\gamma_i > 0$, $i = \overline{1,n}$, где G: $[0; 1]^n \to (0; 1]$ - гладкая строго монотонно возрастающая по каждой из переменных *функция агрегирования*, значение которой $g(t) = G(X(t))$ может интерпретироваться как уровень научения системы в целом – *агрегированный уровень научения*.

В силу системы (17) можно записать следующее уравнение динамики агрегированного уровня научения:

(20) $\dot{g}(t) = \sum_{i=1}^{n} \frac{\partial G(X)}{\partial x_i} \dot{x}_i(t) = g(t) \sum_{i=1}^{n} \frac{\partial G(X)}{\partial x_i} \gamma_i (1 - x_i)$.

В рамках введенных предположений правая часть выражения (20) строго положительна, поэтому агрегированный уровень научения монотонно возрастает, асимптотически стремясь к единице.

Подставляя $f_i(X) = \gamma_i G(X)$ в (17), получим

(21) $\frac{\dot{x}_i(t)}{\gamma_i (1 - x_i)} = G(X), \ i = \overline{1,n}$.

При большом числе агентов можно в первом приближении (т.н. приближение *среднего поля*) пренебречь влиянием отдельного агента на значение агрегированного уровня научения. Тогда из (21) получаем, что справедлива оценка:

(22) $x_i^{\sim}(t) = 1 - (1 - x_i(0)) \exp(-\gamma_i \int_0^t g(\tau) d\tau)$.

Рассмотрим случай $G(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$. Выражение (20) примет вид:

(23) $\dot{g}(t) = g(t) \frac{1}{n} \sum_{i=1}^{n} \gamma_i (1 - x_i(0)) \exp(-\gamma_i \int_0^t g(\tau) d\tau)$.

При больших n в силу результатов [2] среднее в правой части (23) может быть оценено как $\gamma^{\sim}(1 - g(t))$, где $\gamma^{\sim} = \frac{1}{n}\sum_{i=1}^{n} \gamma_i$. Получаем следующее уравнение динамики агрегированного уровня научения:

(24) $\dot{g}(t) = \gamma^{\sim} g(t)(1 - g(t))$

с начальным условием $g(0) = \frac{1}{n}\sum_{i=1}^{n} x_i(0)$. Решением дифференциального уравнения Бернулли (24) является логистическая кривая

(25) $g(t) = \frac{1}{1 + (\frac{1}{g(0)} - 1) \exp(-\gamma^{\sim} t)}$.

Вычислим $\int_0^t g(\tau) d\tau = \frac{1}{\gamma^{\sim}} \ln[g(0)(\exp(\gamma^{\sim} t) + \frac{1}{g(0)} - 1)]$ и подставим в выражение (22):

(26) $x_i^{\sim}(t) = 1 - \dfrac{1 - x_i(0)}{[1 + g(0)(\exp(\gamma^{\sim} t) - 1)]^{\frac{\gamma_i}{\gamma^{\sim}}}}$.

Отметим, что при $n = 1$ выражения (26) и (25) совпадают. Величина $\int_0^t g(\tau) d\tau$, которая асимптотически линейна по t, может интерпретироваться как «эффективное время» в обучаемой системе.

8. Сетевая модель научения

Рассмотрим конечное множество $N = \{1, 2, \ldots, n\}$ *агентов*, $n \geq 2$, и *сеть* $G = (N, E)$ (ориентированный связный граф без циклов), вершины которой соответствуют агентам, а множество дуг $E \subseteq N \times N$ отражает «технологические» связи между агентами, причем номера агентов образуют правильную нумерацию вершин сети. Обозначим через $N_i = \{ j \in N \mid (j; i) \in E \}$ множество *предшественников* i-го агента в сети G, $i \in N$.

Предположим, что сеть имеет единственный *выход* (вершину, не имеющую исходящих дуг) – n-ю вершину. Обозначим через $M_0 \subseteq N$ множество *входов* рассматриваемой сети (вершин, не имеющих входящих дуг), через M_k – множество вершин, в которые входят дуги только из вершин, принадлежащих множествам $\{M_j\}$, $j = \overline{0,k-1}$ (число $k(i)$ называется *рангом* вершины i, принадлежащей множеству M_k), $k = \overline{1,m}$, $m \leq n - 1$, $M_m = \{n\}$. Набор множеств $\{M_k\}$, $k = \overline{0,m}$, является разбиением множества N. Обозначим через $M^k = \bigcup_{j=0}^{k-1} M_k$, $k = \overline{1,m}$, и положим $M^0 = \varnothing$.

Будем считать, что направленные связи между агентами отражают возможности их научения: вероятность освоения агентом «своего» компонента деятельности зависит (равна произведению $\prod_{j \in N_i} x_j(t)$) от уровней научения его предшественников.

Получаем, что сеть структурирует правые части уравнений (17) следующим образом:

(27) $f_i(X) = \gamma_i \prod_{j \in N_i} x_j(t)$, $\gamma_i > 0$, $i = \overline{1,n}$.

Действительно, система уравнений

(28) $\dot{x}_i(t) = \gamma_i (1 - x_i) \prod_{j \in N_i} x_j(t)$, $i = \overline{1,n}$,

с учетом структуры сети допускает последовательное (по множествам $\{M_k\}$, $k = \overline{1,m}$) интегрирование. Входы сети (агенты из множества M_0) будут научаться по экспоненциальному закону (для простоты будем считать, что начальные значения уровней научения всех агентов равны нулю):

(29) $x_i(t) = 1 - \exp(-\gamma_i t)$, $i \in M_0$.

Подставляя (29) в (27), для агентов из множества M_1 получаем:

(30) $\dot{x}_i(t) = \gamma_i (1 - x_i) \prod_{j \in M_0} \left(1 - \exp(-\gamma_j t)\right)$, $i \in M_1$.

Решая (30), запишем:

(31) $x_i(t) = 1 - \exp\{-\gamma_i \int_0^t \prod_{j \in M_0} \left(1 - \exp(-\gamma_j \tau)\right) d\tau$, $i \in M_1$.

И так далее, в общем случае для агентов из множества M_k получаем:

(32) $x_i(t) = 1 - \exp\{-\gamma_i \int_0^t \prod_{j \in M_{k-1}} x_j(\tau) d\tau,\ i \in M_k, k = \overline{1,m}.$

Пример. Пусть сеть имеет вид цепочки из двух агентов, упорядоченных в соответствии со своими номерами. Тогда из (31) получаем, что $M_0 = \{1\}$, $M_1 = \{2\}$ и $x_1(t) = 1 - \exp(-\gamma_1 t)$,

$$x_2(t) = 1 - \exp\{-\gamma_2 \int_0^t \left(1 - \exp(-\gamma_1 \tau)\right) d\tau = 1 - \exp(-\gamma_2 t) \exp(\frac{\gamma_2}{\gamma_1} x_1(t)).$$ •

9. Стохастические оценки

До сих пор мы рассматривали уровень научения как вероятность того, что обучаемый встретится с уже известной ситуацией. Рассмотрим теперь модель, в которой процесс научения описывается совокупностью ситуаций, деятельность в условиях которых им уже освоена (базовая модель рассмотрена в [1]).

Исследуем случай равномерного распределения вероятностей реализации различных ситуаций: $p_k = \frac{1}{K}$. В этом случае процесс обучения может быть описан простой марковской цепью с числом состояний $K + 1$. Номер состояния соответствует количеству ситуаций, для которых соответствующие компоненты деятельности уже освоены. Из состояния n возможен либо возврат в состояние n либо переход в состояние $n + 1$ – см. Рис. 2. Обозначим долю освоенных ситуаций в момент времени t через L_t.

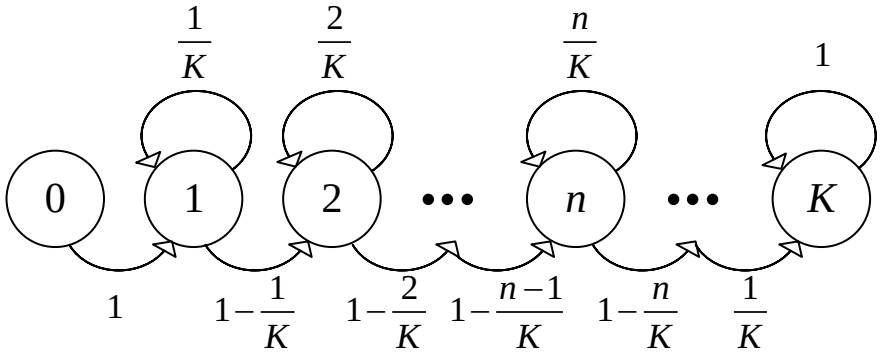

Рис. 2. Диаграмма марковской цепи процесса обучения для случая равномерного распределения

Обозначим для упрощения выкладок $\delta = 1 / K$. Переходные вероятности принимают значения

$$\Pr\left(L_t = n\delta \mid L_{t-1} = (n-1)\delta\right) = 1 - (n-1)\delta,$$
$$\Pr\left(L_t = n\delta \mid L_{t-1} = n\delta\right) = n\delta.$$

Уравнение динамики принимает вид

(32) $\Pr\left(L_t = n\delta\right) = \Pr\left(L_{t-1} = n\delta\right) n\delta + \Pr\left(L_{t-1} = (n-1)\delta\right)\left(1 - (n-1)\delta\right)$.

Поставим задачу найти в явном виде выражение для $\Pr\left(L_t = n\delta\right)$ при $\Pr\left(L_0 = 0\right) = 1$. При $n = 0$, $n = 1$, очевидно, $\Pr\left(L_t = 0\right) = \chi_{t=0}\left(t\right)$, где χ_A – индикатор множества A.

(33) $\Pr\left(L_t = 1\right) = \delta^{t-1}$, $t \geq 1$.

При $t = n$, $n > 1$, имеем

(34) $\Pr\left(L_n = n\delta\right) = \left(1-\delta\right)\left(1-2\delta\right)...\left(1-\left(n-1\right)\delta\right) = \prod_{i=1}^{n-1}\left(1-i\delta\right)$.

При $n = 2$ из (32) следует

(35) $\Pr\left(L_t = 2\delta\right) = \Pr\left(L_{t-1} = 2\delta\right)2\delta + \Pr\left(L_{t-1} = \delta\right)\left(1-\delta\right) = 2\delta\Pr\left(L_{t-1} = 2\delta\right) + \delta^{t-2}\left(1-\delta\right)$.

Выражая $\Pr\left(L_{t-1} = 2\delta\right)$ по формуле (35) и подставляя в правую часть (35), получаем

$$\Pr\left(L_t = 2\delta\right) = 2\delta\left(2\delta\Pr\left(L_{t-2} = 2\delta\right) + \delta^{t-3}\left(1-\delta\right)\right) + \delta^{t-2}\left(1-\delta\right).$$

Продолжая выражать $\Pr\left(L_{t-k} = 2\delta\right)$ через $\Pr\left(L_{t-k-1} = 2\delta\right)$ и раскрывая скобки, имеем

(36) $\Pr\left(L_t = 2\delta\right) = \left(1-\delta\right)\delta^{t-2}\sum_{i=0}^{i=k-1} 2^i + 2^k\delta^k\Pr\left(L_{t-k} = 2\delta\right) = \left(2^k - 1\right)\left(1-\delta\right)\delta^{t-2} + 2^k\delta^k\Pr\left(L_{t-k} = 2\delta\right)$.

Подставим в (36) значение $t - k = 2$. Учитывая $\Pr\left(L_2 = 2\delta\right) = \left(1-\delta\right)$, получаем

$$\Pr\left(L_t = 2\delta\right) = \left(2^{t-2} - 1\right)\left(1-\delta\right)\delta^{t-2} + 2^{t-2}\delta^{t-2}\left(1-\delta\right) = \left(2^{t-1} - 1\right)\delta^{t-2}\left(1-\delta\right).$$

Итак,

(37) $\Pr\left(L_t = 2\delta\right) = \left(2^{t-1} - 1\right)\delta^{t-2}\left(1-\delta\right),\ t \geq 2$

Вид формулы (37) и её вывод (а также выражение (34) для $t = n$) позволяют выдвинуть гипотезу о том, что вероятность $\Pr\left(L_t = n\delta\right)$ имеет вид

(38) $\Pr\left(L_t = n\delta\right) = \left(\alpha_n^n n^{t-1} + \alpha_{n-1}^n\left(n-1\right)^{t-1} + ... + \alpha_1^n\right)\delta^{t-n}\left(1-\delta\right)...\left(1-\left(n-1\right)\delta\right)$.

где $t \geq n$, а $\{\alpha_i^n\}$ – некоторые коэффициенты, не зависящие от t. Для удобства введём обозначения $a_t^n = \sum_{i=1}^{n}\alpha_i^n i^{t-1}$, $b_t^n = \delta^{t-n}\prod_{i=1}^{n-1}\left(1-i\delta\right)$. Тогда гипотеза (38) принимает следующий вид.

Лемма 2. $\Pr\left(L_t = n\delta\right) = a_t^n b_t^n,\ \ t \geq n.$

Докажем Лемму 2 по индукции. При $n = 2$ она верна. Предположим, что утверждение леммы выполняется для $n-1$. Докажем, что тогда она верно для n. Запишем уравнение динамики (32):

(39) $\Pr\left(L_t = n\delta\right) = n\delta\Pr\left(L_{t-1} = n\delta\right) + \left(1-\left(n-1\right)\delta\right)a_{t-1}^{n-1}b_{t-1}^{n-1}$.

Выражая $\Pr\left(L_{t-1} = n\delta\right)$ из (39) и подставляя в правую часть (39), получаем

(40) $\Pr\left(L_t = n\delta\right) = n\delta\left(n\delta\Pr\left(L_{t-2} = n\delta\right) + \left(1-\left(n-1\right)\delta\right)a_{t-2}^{n-1}b_{t-2}^{n-1}\right) + \left(1-\left(n-1\right)\delta\right)a_{t-1}^{n-1}b_{t-1}^{n-1}$.

Продолжая выражать $\Pr\left(L_{t-k} = n\delta\right)$ через $\Pr\left(L_{t-k-1} = n\delta\right)$ и раскрывая скобки, а также принимая во внимание, что $\delta b_{t-1}^n = b_t^n$, $\left(1-\left(n-1\right)\delta\right)b_t^{n-1} = \delta b_t^n$, имеем

$$\Pr\left(L_t = n\delta\right) = n^k\delta^k\Pr\left(L_{t-k} = n\delta\right) + b_t^n\sum_{i=0}^{i=k-1} n^i a_{t-1-i}^{n-1}.$$

Подставляя $k = t - n$, с учётом $a_{n-1}^{n-1} = 1$ (что следует из (34)), получаем

$$\Pr\left(L_t = n\delta\right) = n^{t-n}b_t^n + b_t^n\sum_{j=0}^{j=t-n-1} n^j a_{t-j-1}^{n-1} = b_t^n\sum_{j=0}^{j=t-n} n^j a_{t-j-1}^{n-1} = b_t^n\sum_{j=0}^{j=t-n} n^j\sum_{i=1}^{n-1}\alpha_i^{n-1}i^{t-j-2} =$$

$$= b_t^n\sum_{i=1}^{n-1}\alpha_i^{n-1}\sum_{j=0}^{j=t-n} n^j i^{t-j-2} = b_t^n\sum_{i=1}^{n-1}\alpha_i^{n-1}i^{n-2}\sum_{j=0}^{j=t-n} n^j i^{t-n-j} = b_t^n\sum_{i=1}^{n-1}\alpha_i^{n-1}i^{n-2}\frac{n^{t-n+1} - i^{t-n+1}}{n-i} =$$

$$= b_t^n\sum_{i=1}^{n-1}\left(-\frac{\alpha_i^{n-1}}{n-i}\right)i^{t-1} + n^{t-1}\left(n^{2-n}\sum_{i=1}^{n-1}\frac{\alpha_i^{n-1}i^{n-2}}{n-i}\right) = b_t^n\sum_{i=1}^{n}\alpha_i^n i^{t-1} = a_t^n b_t^n,$$

где

(41) $\alpha_i^n = -\dfrac{\alpha_i^{n-1}}{n-i}$, $i < n$,

(42) $\alpha_n^n = n^{2-n}\sum\limits_{i=1}^{n-1}\dfrac{\alpha_i^{n-1} i^{n-2}}{n-i} = -n^{2-n}\sum\limits_{i=1}^{n-1}\alpha_i^n i^{n-2}$.

Лемма 2 доказана.

Для нахождения явного вида вероятности $\Pr(L_t = n\delta)$ остаётся определить явный вид коэффициентов α_i^n. Используя формулу (41), получаем

(43) $\alpha_i^n = -\dfrac{\alpha_i^{n-1}}{n-i} = \dfrac{\alpha_i^{n-2}}{(n-i)(n-1-i)} = \ldots = \dfrac{(-1)^{n+i}}{(n-i)!}\alpha_i^i$.

Докажем, что

(44) $\alpha_i^i = \dfrac{1}{(i-1)!}$

по индукции. Для $i = 1$, $i = 2$ выражение (44) верно, согласно (33) и (37). Пусть (44) верно для всех i от 1 до $n - 1$. Докажем, что тогда оно верно и для $i = n$.

Подставляя в (42) выражения (42), (44), получаем

$$\alpha_n^n = -n^{2-n}\sum_{i=1}^{n-1}\alpha_i^n i^{n-2} = -n^{2-n}\sum_{i=1}^{n-1}\frac{(-1)^{n+i}}{(n-i)!}\alpha_i^i i^{n-2} = -n^{2-n}\sum_{i=1}^{n-1}\frac{(-1)^{n+i}}{(n-i)!(i-1)!}i^{n-2} =$$

$$= \frac{1}{(n-1)!} - n^{2-n}(-1)^n\sum_{i=1}^{n}\frac{(-1)^i}{(n-i)!(i-1)!}i^{n-2} = \frac{1}{(n-1)!} - \frac{n^{2-n}(-1)^n}{n!}\sum_{i=1}^{n}(-1)^i C_n^i i^{n-1}.$$

Известно (см. например [3]), что $\sum\limits_{i=1}^{n}(-1)^i C_n^i i^{n-1} = 0$. Следовательно, $\alpha_n^n = \dfrac{1}{(n-1)!}$.

Объединяя (43) и (44), получаем, что

(45) $\alpha_i^n = \dfrac{(-1)^{n+i}}{(n-i)!(i-1)!}$, $i < n$.

Подставляя (45) в (38), получаем окончательно:

(46) $\Pr(L_t = n\delta) = \left(\sum\limits_{i=1}^{n}\dfrac{(-1)^{n+i} i^{t-1}}{(n-i)!(i-1)!}\right)\delta^{t-n}(1-\delta)\ldots(1-(n-1)\delta) =$

$$= \left(\sum_{i=1}^{n}(-1)^{n+i}\frac{C_n^i}{n!}i^t\right)\delta^{t-n}(1-\delta)\ldots(1-(n-1)\delta).$$

Используя (46), можно численно строить доверительные интервалы. Например, найти интервал значений (a, b), такой, что $L_t \in (a,b)$ с заданной вероятностью p.

Из вышеизложенного следует справедливость следующего утверждения.

Утверждение 5. Имеют место следующие стохастические оценки процесса научения:

1. $\lim\limits_{t\to\infty}\dfrac{1}{t}\ln\Pr(L_t = n\delta) = -\ln\dfrac{K}{n}$ (большие уклонения);
2. $\lim\limits_{t\to\infty}\dfrac{1}{t}\ln\Pr(L_t < 1) = -\ln\dfrac{K}{K-1}$ (большие уклонения);
3. $\Pr(L_t = 1) = \left(\sum\limits_{i=1}^{n}(-1)^{K+i}\dfrac{C_K^i}{K!}i^t\right)\dfrac{1}{K^{t-K}}\left(1-\dfrac{1}{K}\right)\ldots\left(1-\dfrac{(K-1)}{K}\right)$;

4. $\Pr\left(\frac{l}{K} \le L_t \le \frac{m}{K}\right) = \sum_{n=l}^{n=m} \sum_{i=1}^{n} (-1)^{n+i} \frac{C_n^i}{n!} i^t \delta^{t-n} \prod_{i=1}^{n-1} (1 - i\delta)$.

Литература

1 Белов М.В., Новиков Д.А. Модели технологий. – М.: Ленанд, 2019. – 160 с.

2 Опойцев В.И. Задачи и проблемы асимптотического агрегирования // Автоматика и телемеханика. 1991. N 8. С. 133 – 144.

3 Gould H. Euler's Formula for Nth Differences of Powers // The American Mathematical Monthly. 1978. Vol. 85. N. 6. - P. 450 – 467.

4 Lasry J., Lions P. Mean Field Games // Jpn. J. Math. 2007. Vol. 2. N 1. P. 229 – 260.