О марковских случайных полях и их связи с цепями Маркова

А. А. Петюшко

В данной работе рассматриваются марковские и скрытые марковские случайные поля. Устанавливаются взаимосвязи между марковскими случайными полями и цепями Маркова, а также между скрытыми марковскими полями и скрытыми марковскими моделями посредством введенных порожденных случайных величин.

Ключевые слова: марковское случайное поле, скрытое марковское случайное поле, марковская цепь, скрытая марковская модель.

1. Необходимые определения

Пусть $A = \{1, 2, \ldots, a\}$ и $B = \{1, 2, \ldots b\}$, $a, b < \infty$ — два конечных множества. Пусть $S = \{1, 2, \ldots, N\}$ — конечное множество индексов.

Пусть $X = \{X_i | \; i \in S\}$ — многомерная случайная величина, такая что каждая компонента X_j, являющаяся одномерной случайной величиной, принимает значение x_j и определена в своем вероятностном пространстве. Считается, что $\forall j \; X_j$ дискретны, определены на одном вероятностном пространстве и множество значений — конечно.

Для удобства рассмотрения можно представлять, что множество индексов S задает множество точек на плоскости. Соответственно, рассматривается реализация многомерной случайной величины X в этих точках. Введенная таким образом случайная величина X называется *случайным полем* (сокращенно СП; в англоязычной литературе принято название Random Field).

Конкретная реализация $x = (x_1, \ldots, x_N)$ многомерной случайной величины X, то есть совместное событие $(X_1 = x_1, \ldots, X_N = x_N)$ (кратко $X = x$), называется *конфигурацией* X.

Пусть X — случайное поле со значениями на множестве A, то есть $\forall i \; X_i \in A$. Если x — какая-то конкретная конфигурация X, то χ — множество всех возможных конфигураций:

$$\chi = \{x = (x_1, \ldots, x_N) | \; x_i \in A \; \forall i \in S\}.$$

Система соседства — это множество $\partial = \{\partial i | \; i \in S\}$, где ∂i — множество элементов из S, называемое *шаблоном соседства для элемента* i, такое что:

$$\begin{cases} i \notin \partial i, \\ i \in \partial j \Leftrightarrow j \in \partial i. \end{cases}$$

Например, на рис. 1 система соседства $\partial = \{\{i-1, i+1\} | \; i \in S\}$.

Определение. Случайное поле X называется *марковским случайным полем* (сокращенно МСП; в англоязычной литературе Markov Random Field, MRF [1]) в соответствии с системой соседства ∂ тогда и только тогда, когда $\forall i$:

$$\begin{cases} P(X = x) > 0 \; \forall x \in \chi, \\ P(X_i = x_i | \; X_j = x_j, \; j \in S \backslash \{i\}) = P(X_i = x_i | \; X_j = x_j, \; j \in \partial i). \end{cases}$$

Примеры марковских случайных полей с различными вариантами условной локальной зависимости можно посмотреть на рис. 1, 2, 3 и 4.

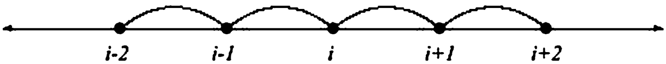

Рис. 1. Линейное марковское случайное поле с условной локальной зависимостью для элемента i вида $P(X_i | X_{i-1}, X_{i+1})$.

Определение. *Клика* c для системы соседства ∂ — множество элементов из S, такое что $\forall s, r \in c, s \neq r \Rightarrow r \in \partial s$. Несложно заметить, что любое подмножество клики — также клика (считается, что любое одноэлементное множество также является кликой; в этом случае

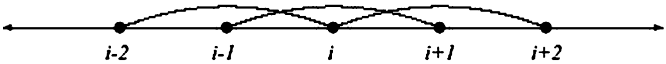

Рис. 2. Линейное марковское случайное поле с условной локальной зависимостью для элемента i вида $P(X_i|X_{i-2}, X_{i+2})$.

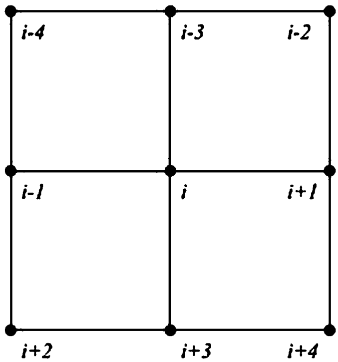

Рис. 3. Марковское случайное поле с условной локальной зависимостью для элемента i типа «крест» $P(X_i|X_{i-3}, X_{i-1}, X_{i+1}, X_{i+3})$.

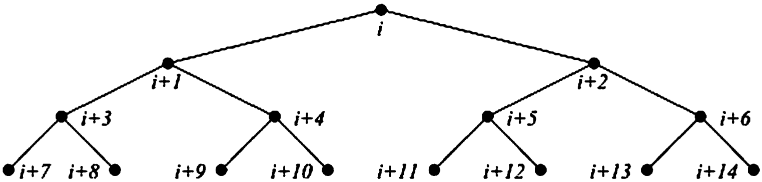

Рис. 4. Марковское случайное поле, заданное на дереве, с условной локальной зависимостью для элемента $i+1$ вида $P(X_{i+1}|X_i, X_{i+3}, X_{i+4})$.

такая клика называется тривиальной). Например, кликами, содержащими элемент i (рис. 1), будут являться множества $\{i, i+1\}$, $\{i-1, i\}$ и тривиальная клика $\{i\}$.

Пусть c — клика, а x_c — ограничение конфигурации x на c, то есть $x_c = (x_{i_1}, x_{i_2}, \ldots, x_{i_{|c|}})$, где $i_j \in c$, $j = 1 \ldots |c|$. Пусть $C(\partial)$ —

множество всех клик для системы соседства ∂, тогда *потенциальная функция* $V_c(x_c)$ определяется как любая функция $V : C(\partial) \to R$.

Определение. Дискретное распределение называется *распределением Гиббса* [1], если

$$\mathbf{P}(X = x) = \frac{1}{Z} \exp\left(- \sum_{c \in C(\partial)} V_c(x_c) \right), \tag{1}$$

где Z — нормирующая константа, такая что:

$$Z = \sum_{x \in \chi} \exp\left(- \sum_{c \in C(\partial)} V_c(x_c) \right). \tag{2}$$

Очевидно, что в случае нулевых потенциальных функций для какого-нибудь набора клик, соответствующие члены будут просто отсутствовать в формулах (1), (2).

Наиболее важной теоремой, связывающей марковские случайные поля и распределение Гиббса, является следующая

Теорема 1 (Hammersley-Clifford [2]). *X — марковское случайное поле тогда и только тогда, когда $\mathbf{P}(X = x)$ — распределение Гиббса.*

Таким образом, имеется возможность вычислять вероятность конфигурации для любого марковского случайного поля по формулам (1), (2).

Определение. *Скрытое марковское случайное поле* (сокращенно СМСП; в англоязычной литературе Hidden Markov Random Field, HMRF [3]) — пара случайный полей (X, Y), такая что:

1) $X = \{X_i |\ i \in S\}$ — так называемое «скрытое» (или, другими словами, ненаблюдаемое) марковское поле со значениями в A.
2) $Y = \{Y_i |\ i \in S\}$ — наблюдаемое (вовсе не обязательно марковское) случайное поле со значениями в B. Важно, что $\forall i \in S$, $\forall d \in A$ известны условные распределения $\mathbf{P}(Y_i |\ X_i = d)$.

3) Для любой конфигурации $x \in \chi$ случайные величины Y_i условно независимы, то есть
$$\mathbf{P}(Y|\ X = x) = \prod_{i \in S} \mathbf{P}(Y_i|\ X_i = x_i).$$

Скрытое марковское случайное поле можно себе представлять как наблюдаемое случайное поле Y, каждой точке которого соответствует ненаблюдаемое (скрытое) марковское случайное поле X. Например, скрытое марковское случайное поле, соответствующее марковскому случайному полю на рис. 1, изображено на рис. 5.

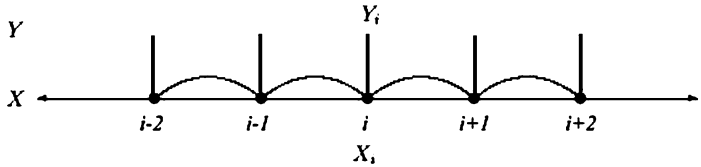

Рис. 5. Скрытое марковское случайное поле, соответствующее марковскому случайному полю на рис. 1.

Очевидно, что
$$\mathbf{P}(Y = y, X = x) = \mathbf{P}(Y = y|\ X = x)\mathbf{P}(X = x) = \\ = \mathbf{P}(X = x) \prod_{i \in S} \mathbf{P}(Y_i = y_i|\ X_i = x_i).$$

Совместное распределение (X_i, Y_i) с шаблоном соседства ∂i определяется как
$$\mathbf{P}(X_i = x_i, Y_i = y_i|\ X_j = x_j, j \in \partial i) = \\ = \mathbf{P}(Y_i = y_i|\ X_i = x_i)\mathbf{P}(X_i = x_i|\ X_j = x_j, j \in \partial i),$$
а распределение Y_i с шаблоном соседства ∂i определяется как
$$\mathbf{P}(Y_i = y_i|\ X_j = x_j, j \in \partial i) = \\ = \sum_{d \in A} \mathbf{P}(Y_i = y_i, X_i = d|\ X_j = x_j, j \in \partial i) = \\ = \sum_{d \in A} \mathbf{P}(Y_i = y_i|\ X_i = d)\mathbf{P}(X_i = d|\ X_j = x_j, j \in \partial i).$$

2. Взаимосвязи между МСП (СМСП) и марковскими цепями (скрытыми марковскими моделями)

Рассмотрим марковскую цепь (сокращенно МЦ) с состояниями $x_i \in A$, функционирующую в дискретном времени. Построим многомерную случайную величину $X = \{X_i |\ i = \overline{0 \dots n}\}$ по этой марковской цепи следующим образом: $\mathbf{P}(X_i = x_i)$ — это вероятность того, что в момент времени i мы находились в состоянии x_i. Назовем X *порожденной* случайной величиной (сокращенно п.с.в.).

Утверждение 1. *Любая порожденная случайная величина — это марковское случайное поле линейной структуры.*

Доказательство. Зададим шаблон соседства для $i : \partial i = \{i-1, i+1\}$. Любая нетривиальная клика c будет иметь вид $c = \{i-1, i\}$ (для любой тривиальной клики будем полагать соответствующую ей потенциальную функцию нулевой). Тогда вероятность реализации порожденной случайной величины X (то есть прохода по марковской цепи по соответствующим состояниям) будет иметь вид

$$\mathbf{P}(X = x) = \mathbf{P}(X_0 = x_0) \prod_{i=1}^{n} \mathbf{P}(X_i = x_i |\ X_{i-1} = x_{i-1}) =$$
$$= \mathbf{P}(X_0 = x_0) \exp\left(\sum_{i=1}^{n} \log \mathbf{P}(X_i = x_i |\ X_{i-1} = x_{i-1}) \right),$$

то есть будет иметь вид распределения Гиббса с потенциальными функциями

$$V(x_i, x_{i-1}) = -\log \mathbf{P}(X_i = x_i |\ X_{i-1} = x_{i-1}).$$

Утверждение доказано.

Утверждение 2. *Любое марковское случайное поле линейной структуры — это порожденная случайная величина.*

Доказательство. Исходя из определения случайного марковского поля, можно записать вероятность нахождения в данном состоянии в зависимости от нахождений во всех иных состояниях:

$$\mathbf{P}(X_i = x_i|\ X_j = x_j, j \neq i) = \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}).$$

Зависимость от состояний с меньшими номерами выражается формулой

$$\mathbf{P}(X_i = x_i|\ X_j = x_j, j < i) = \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}).$$

Таким образом, зависимость от предыстории и от всей последовательности состояний у линейного марковского случайного поля такие же, как и у порожденной от марковской цепи случайной величины. Утверждение доказано.

Отметим, что в обоих утверждениях мы устанавливали связь с марковским случайным полем, которое изображено на рис. 1. Также обратим внимание на то, что вычисление вероятностей перехода по марковскому случайному полю не является тривиальной задачей и в данной работе не рассматривается.

Теперь рассмотрим скрытые марковские модели [4] и их связь с марковскими случайными полями и скрытыми марковскими случайными полями.

Определение. *Скрытая марковская модель* (сокращенно СММ; в англоязычной литературе Hidden Markov Model, HMM) — это дважды случайный процесс, то есть модель, состоящая из N состояний, в каждом из которых некоторая система может принимать одно из M значений какого-либо параметра. Вероятности переходов между состояниями задается матрицей вероятностей $A = (a_{ij})$, где a_{ij} — вероятность перехода из i-го в j-е состояние. Вероятности выпадения каждого из M значений параметра в каждом из N состояний задается матрицей $B = (b_{jk})$, где (b_{jk}) — вероятность выпадения k-го значения параметра в j-м состоянии. Вероятность наступления начального состояния задается вектором $\pi = (\pi_i)$, где π_i — вероятность того, что в начальный момент система окажется в i-м состоянии.

Полагается, что наблюдаемые значения Y_k зависят только от соответствующего им скрытого состояния марковской цепи X_j, то есть $\mathbf{P}(Y = y|\ X = x) = \prod_{i=0}^{n} \mathbf{P}(Y_i = y_i|\ X_i = x_i)$. Таким образом, скрытой марковской моделью называется тройка $\lambda = (A, B, \pi)$.

Возьмем скрытую марковскую модель, функционирующую в дискретном времени. Пусть при некотором ее функционировании при проходе через состояния $x_0, x_1, \ldots, x_n$ на выход подавались буквы $y_0, y_1, \ldots, y_n$ соответственно. Рассмотрим многомерную случайную величину

$$Z = (Z_1, Z_2, \ldots, Z_{2n+2}) = (X_0, Y_0, X_1, Y_1, \ldots, X_n, Y_n),$$

причем вероятность $\mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1})$ — это вероятность перехода $a(x_{i-1}, x_i)$ из состояния x_{i-1} в состояние x_i, а вероятность $\mathbf{P}(Y_i = y_i|\ X_i = x_i)$ — это вероятность $b(y_i, x_i)$ выдачи буквы y_i в состоянии x_i. Назовем такую многомерную случайную величину *скрытой порожденной* случайной величиной (сокращенно с.п.с.в.).

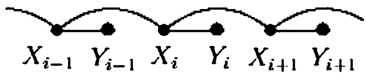

Рис. 6. Иллюстрация связи между скрытой порожденной случайной величиной и марковским случайным полем.

Утверждение 3. *Любая скрытая порожденная случайная величина — это случайное марковское поле специального вида (см. рис. 6).*

Доказательство. Для скрытой порожденной случайной величины нетривиальные клики, содержащие номер $2k+1$: $c_1 = \{2k-1, 2k+1\}$, $c_2 = \{2k+1, 2k+2\}, c_3 = \{2k+1, 2k+3\}$ (по сути, клики c_1 и c_3 идентичны); нетривиальная клика, содержащая номер $2k+2$ — это c_2. Вероятность реализации скрытой порожденной случайной величины Z есть

$$\mathbf{P}(Z = z) = \mathbf{P}(X = x, Y = y) = \mathbf{P}(X_0 = x_0)\mathbf{P}(Y_0 = y_0|\ X_0 = x_0) \times$$

$$\times \prod_{i=1}^{n} \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1})\mathbf{P}(Y_i = y_i|\ X_i = x_i) =$$

$$= \mathbf{P}(X_0 = x_0) \exp\Bigg(\log \mathbf{P}(Y_0 = y_0|\ X_0 = x_0) +$$

$$+ \sum_{i=1}^{n} (\log \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}) + \log \mathbf{P}(Y_i = y_i|\ X_i = x_i)) \Bigg) =$$

$$= \mathbf{P}(X_0 = x_0) \exp\left(- \sum_{i=1}^{n} V_{c_1}(x_i, x_{i-1}) - \sum_{i=0}^{n} V_{c_2}(y_i, x_i) \right),$$

то есть вероятность реализации задается распределением Гиббса с потенциальными функциями

$$V_{c_1}(x_i, x_{i-1}) = -\log \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}) = -\log a(x_{i-1}, x_i)$$

и

$$V_{c_2}(y_i, x_i) = -\log \mathbf{P}(Y_i = y_i|\ X_i = x_i) = -\log b(y_i, x_i).$$

Таким образом, по теореме Хаммерсли-Клиффорда скрытая порожденная случайная величина является случайным марковским полем. Утверждение доказано.

Утверждение 4. *Любая скрытая порожденная случайная величина* $Z = (Z_1, Z_2, \ldots, Z_{2n+2}) = (X_0, Y_0, X_1, Y_1, \ldots, X_n, Y_n)$ *— это скрытое случайное марковское поле* (X, Y) *линейной структуры.*

Доказательство. Для доказательства утверждения установим, что все три пункта в определении скрытого случайного марковского поля выполняются.

Доказательство того, что X — случайное марковское поле с потенциальными функциями вида $V(x_i, x_{i-1}) = -\log \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}) = -\log a(x_{i-1}, x_i)$ проходит совершенно аналогично доказательству предыдущего утверждения. Условные вероятности $\mathbf{P}(Y_i = y_i|\ X_i = x_i)$ также даны и равны $b(y_i, x_i)$. И, наконец, по определению скрытой марковской модели Y_i условно независимы, так как наблюдаемая буква зависит только от скрытого под ней состояния:

$$\mathbf{P}(Y = y|\ X = x) = \prod_{i=0}^{n} \mathbf{P}(Y_i = y_i|\ X_i = x_i).$$

Таким образом, скрытая порожденная случайная величина является скрытым марковским случайным полем линейной структуры. Утверждение доказано.

Утверждение 5. *Любое скрытое марковское случайное поле линейной структуры — это скрытая порожденная случайная величина.*

Доказательство. Имеется скрытое марковское случайное поле с наблюдаемым случайным полем $Y = \{Y_0, Y_1, \dots, Y_n\}$ и скрытым марковским случайным полем $X = \{X_0, X_1, \dots, X_n\}$.

Рассмотрим многомерную случайную величину

$$Z = (Z_1, Z_2, \dots, Z_{2n+2}) = (X_0, Y_0, X_1, Y_1, \dots, X_n, Y_n).$$

Исходя из определения скрытого случайного марковского поля, можно записать вероятность нахождения в данном состоянии в зависимости от нахождений во всех иных состояниях:

$$\mathbf{P}(X_i = x_i|\ X_j = x_j, j \neq i) = \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}).$$

Зависимость от состояний с меньшими номерами выражается формулой

$$\mathbf{P}(X_i = x_i|\ X_j = x_j, j < i) = \mathbf{P}(X_i = x_i|\ X_{i-1} = x_{i-1}).$$

Таким образом, случайная величина $X = \{X_0, X_1, \dots, X_n\}$ — порожденная случайная величина (как и должно быть, так как она должна быть порождена от ненаблюдаемой марковской цепи). А случайная величина Z является скрытой порожденной случайной величиной, так как по определению скрытого случайного марковского поля для любой конфигурации $x \in \chi$ случайные величины Y_i условно независимы:

$$\mathbf{P}(Y|\ X = x) = \prod_{i \in S} \mathbf{P}(Y_i|\ X_i = x_i),$$

то есть наблюдаемая буква зависит только от скрытого под ней состояния. Утверждение доказано.

Отметим, что в двух предыдущих утверждениях мы устанавливали связь со скрытым марковским случайным полем, которое изображено на рис. 5.

$$\begin{array}{lcl} \text{МСП} & \overset{\text{п.с.в.}}{\supset} & \text{МЦ} \\ \cup\ \text{с.п.с.в.} & & \\ \text{СММ} & \underset{\text{с.п.с.в.}}{\subset} & \text{СМСП} \end{array}$$

Рис. 7. Диаграмма вложенности.

В итоге справедлива следующая

Теорема 2. *Имеет место диаграмма вложенности, изображенная на рис. 7.*

Доказательство напрямую следует из утверждений 1–5.

Список литературы

[1] Kindermann R., Snell J. L. Markov Random Fields and Their Applications. American Mathematical Society, 1980.

[2] Hammersley J. M., Clifford P. Markov random fields in statistics. Unpublished paper, 1971.

[3] Zhang Y. Hidden markov random field model. www.fmrib.ox.ac.uk/analysis/techrep/tr00yz1/tr00yz1/node5.html. 2000.

[4] Rabiner L. R. A tutorial on hidden markov models and selected applications in speech recognition // Proceedings of the IEEE. 77 (2). P. 257-–286. February 1989.