Цепи Маркова

Давайте вспомним пример из второй лекции про клиента магазина, перемещающегося по отделам в какой-то последовательности. Это случайный процесс с дискретным временем (т.к. нас интересует только последовательность переходов между отделами, а не время, проведенное в каждом отделе), значения которого лежат в конечном множестве (множество отделов). Этот процесс можно описать с помощью цепи Маркова. Давайте разберемся, что это такое.

4.1 Определение цепи Маркова

Пусть $(X_t)_{t\in\mathbb{Z}_+}$ - случайный процесс на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, описывающий поведение клиента магазина. В данном случаи моменты времени t принимают дискретные значение в $\mathbb{Z}_+ = \{0, 1, 2, ...\}$. Пусть $\Xi = \{1, 2, ...r\}$ - пространство состояний (множество значений, которые могут принимать случайные величины X_t). Т.к. Ξ - дискретное пространство, логичнее всего ввести на нем наибольшую σ-алгебру $\mathcal{F}$, состоящую из всех возможных подмножеств Ξ.

Чтобы определить случайный процесс $(X_t)_{t\in\mathbb{Z}_+}$, нам нужно задать согласованное (т.е. удовлетворяющее теореме Колмогорова) семейство совместных распределений

$$F(t_1, .., t_n; A_1, .., A_n) = P\left(X_{t_1} \in A_1, .., X_{t_n} \in A_n\right), \quad A_1, ..., A_n \in \mathcal{F}.$$

Давайте сделаем это для последовательных моментов времени $t_0 = 0, t_1 = 1, ...t_n = n$ и для одноэлементных множеств вида $A_s = \{i\}$. Пусть задано $\mu \in \mathbb{R}^{1\times r}$ - вектор-строка, описывающая начальное распределение, а также $p_{ij}(t)$ - вероятности перехода из состояния i в состояние j в момент времени t:

$$\mu = (P(X_0 = 1), ...P(X_0 = r)) \in \mathbb{R}^{1\times r},$$
$$p_{ij}(t) = P(X_t = j|X_{t-1} = i).$$

Зададим совместное распределение следующим образом:

$$P\left(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_n = i_n\right) = \mu_{i_0} p_{i_0 i_1}(1) p_{i_1 i_2}(2) ... p_{i_{n-1} i_n}(n). \tag{4.1}$$

В дальнейшем вы должны будем проверить, что эта формула задает семейство согласованных мер.

Упражнение 4.1. *Как в этом случае будут выглядеть совместные распределения для произвольных моментов времени $t_1, t_2, ...t_k$ и произвольных множеств $A_1, A_2, ...A_n \in \mathcal{F}$?*

4.1.1 Стохастические матрицы

Переходные вероятности $p_{ij}(t)$ удобно записать в виде матрицы $P(t) = (p_{ij}(t))_{i,j=1}^{r} \in \mathbb{R}^{r\times r}$. Однако не любая матрица задает семейство переходных вероятностей. Для этого нужно, чтобы элементы матрицы $p_{ij}(t) = P(X_t = j|X_{t-1} = i)$ в самом деле были условными вероятностями.

Упражнение 4.2. *Вспомните, что такое условная вероятность. Докажите, что* $f(A) := P(A|B) = \frac{P(A\cap B)}{P(B)}$ *является вероятностной мерой относительно* A, *т.е.* $0 \le f(A) \le 1$, $f(\Omega) = 1$, $f(\bigsqcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} f(A_i)$.

Таким образом, чтобы матрица $P(t) = (p_{ij}(t))_{i,j=1}^{r} \in \mathbb{R}^{r\times r}$ задавала семейство переходных вероятностей, должны быть выполнены следующие свойства:

1. Элементы матрицы должны быть неотрицательны:

$$p_{ij}(t) = P(X_t = j|X_{t-1} = i) \ge 0,$$

2. Сумма элементов в любой строке должна быть равна единице:

$$\sum_{j=1}^{r} p_{ij}(t) = P(X_t \in \Xi|X_{t-1} = i) = 1.$$

Определение 4.1. *Матрицы, удовлетворяющие выше перечисленным свойствам, называются стохастическими.*

Упражнение 4.3. *Докажите, что если* $Q \in \mathbb{R}^{r\times r}$ *- стохастическая матрица,* $\mu = (\mu_1, ...\mu_r)$ *- распределение вероятностей, то* μQ *- также распределение вероятностей.*

Упражнение 4.4. *Докажите, что если* $Q', Q'' \in \mathbb{R}^{r\times r}$ *- стохастические матрицы, то* $Q := Q'Q''$ *- также стохастическая матрица.*

Определение 4.2. *Таким образом, по теореме Колмогорова начальное распределение* μ *и набор стохастических матриц* $(P(t))_{t\in\mathbb{N}}$ *задают цепь Маркова, определяемую с помощью согласованного семейства совместных распределений по формуле 4.1.*

Упражнение 4.5. *Проверьте, что формула 4.1 задает вероятностную меру на пространстве последовательностей* $\Omega = \{\omega = (i_0, i_1, i_2, ...i_n), i_0, ...i_n \in \Xi\}$, *снабженном наибольшей* σ*-алгеброй* $\mathcal{G} = 2^{\Omega}$, *т.е.*

$$0 \le P(\omega) \le 1, P(\Omega) = 1, P(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \quad \forall A_i \in \mathcal{G},$$

где $\bigsqcup\limits_{i=1}^{\infty} A_i$ *обозначаем объединение непересекающихся множеств* A_i.

Упражнение 4.6. *Почему семейство распределений, заданное формулой 4.1, будет согласованным?*

Теорема 4.1. *Для цепи Маркова, порожденной начальным распределением μ и стохастическими матрицами $P(t)$, выполняются следующие равенства:*

$$P(X_n = j|X_0 = i_0, X_1 = i_1, ...X_{n-1} = i) = p_{ij}(n), \tag{4.2}$$

$$P\left(X_{t+s} = j|X_t = i\right) = \left(P(t+1)...P(t+s)\right)_{ij}, \tag{4.3}$$

$$P\left(X_t = j\right) = \left(\mu P(1)P(2)...P(t)\right)_j. \tag{4.4}$$

Доказательство. Сначала докажем формулу 4.2. Для этого запишем определение условной вероятности $P(A|B) = \frac{P(A\cap B)}{P(B)}$, а после этого используем определение совместного распределения 4.1.

$$P(X_k = j|X_0 = i_0, X_1 = i_1, ...X_{k-1} = i_{k-1}) = \frac{P(X_0 = i_0, X_1 = i_1, ...X_{k-1} = i_{k-1}, X_k = j)}{P(X_0 = i_0, X_1 = i_1, ...X_{k-1} = i_{k-1})} =$$

$$= \frac{\mu_{i_0} p_{i_0 i_1}(1) p_{i_1 i_2}(2) ... p_{i_{n-2} i_{n-1}}(n-1) p_{i_{n-1} i_n}(n)}{\mu_{i_0} p_{i_0 i_1}(1) p_{i_1 i_2}(2) ... p_{i_{n-2} i_{n-1}}(n-1)} = p_{i_{n-1} i_n}(n).$$

Для доказательства формулы 4.3 снова используем определение условной вероятности, после этого применим формулу полной вероятности. А в конце используем матричную запись для стохастических матриц $P(t)$, порождающих цепь Маркова.

$$P\left(X_{t+s} = j|X_t = i\right) = \frac{P(X_{t+s} = j, X_t = i)}{P(X_t = i)} =$$

$$= \frac{\sum\limits_{i_0, i_1, ...i_{t-1}, i_{t+1}, ...i_{t+s-1} \in \Xi} P(X_0 = i_0, X_1 = i_1, ...X_{t-1} = i_{t-1}, X_t = i, ...X_{t+s-1} = i_{t+s-1} X_{t+s} = j)}{\sum\limits_{i_0, i_1, ...i_{t-1} \in \Xi} P(X_0 = i_0, X_1 = i_1, ...X_{t-1} = i_{t-1} X_t = i)} =$$

$$= \frac{\sum\limits_{i_0, i_1, ...i_{t-1} \in \Xi} \mu_{i_0} p_{i_0 i_1}(1) ... p_{i_{t-1} i}(t) \sum\limits_{i_{t+1}, ...i_{t+s-1} \in \Xi} p_{i i_{t+1}}(t+1) ... p_{i_{t+s-1} j}(t+s)}{\sum\limits_{i_0, i_1, ...i_{t-1} \in \Xi} \mu_{i_0} p_{i_0 i_1}(1) ... p_{i_{t-1} i}(t)} =$$

$$= \sum_{i_{t+1}, ...i_{t+s-1} \in \Xi} p_{i i_{t+1}}(t+1) p_{i_{t+1} i_{t+2}}(t+2) ... p_{i_{t+s-2} i_{t+s-1}}(t+s-1) p_{i_{t+s-1} j}(t+s) =$$

$$= \left(P(t+1)P(t+2)...P(t+s)\right)_{ij}.$$

Для доказательства утверждения 4.4 используем формулу полной вероятности, а после этого используем определение совместного распределения 4.1, а также матричную запись для стоххастических матриц $P(t)$, порождающих цепь Маркова.

$$P\left(X_t = j\right) = \sum_{i_0, i_1, \ldots i_{t-1} \in \Xi} P(X_0 = i_0, X_1 = i_1, \ldots X_{t-1} = i_{t-1}, X_t = j) =$$

$$= \sum_{i_0, i_1, \ldots i_{t-1} \in \Xi} \mu_{i_0} p_{i_0 i_1}(1) \ldots p_{i_{t-2} i_{t-1}}(t-1) p_{i_{t-1} j}(t) = \left(\mu P(1) P(2) \ldots P(t)\right)_j .$$

□

4.1.2 Однородная цепь Маркова

Давайте рассмотрим частный случай цепи Маркова, в котором переходные вероятности не зависят от времени. Такая цепь Маркова называется однородной. Эти модели удобны тем, что их можно легко визуализировать с помощью графов. В однородных цепях Маркова выражения $P(X_{t+s} = j | X_t = i)$ и $P(X_t = j)$ принимают более лаконичную форму.

Определение 4.3. *Цепь Маркова называется однородной, если стохастические матрицы, ее определяющие, не зависят от времени, т.е.* $P(t) = P, \forall t \in \mathbb{N}$.

Давайте представим однородную цепь Маркова в виде ориентированного взвешенного графа $G = (V, E, W)$. Множество вершин графа V соответствует пространству состояний $\Xi = \{1, 2, \ldots r\}$ цепи Маркова. Ребро графа $e_{ij} \in E$ имеет вес W_{ij} равный вероятности перехода p_{ij} из состояния i в состояние j. Любая реализация случайного процессса, соответствующего цепи Маркова, представляется как путь в графе. Следовательно, однородную цепь Маркова можно представить как распределение вероятностей на пространстве путей в полученном графе.

Пример 4.1. *Марковской цепи с переходными вероятностями* $P = \begin{pmatrix} 0.1 & 0.9 & 0 \\ 0 & 0.2 & 0.8 \\ 0.7 & 0 & 0.3 \end{pmatrix}$ *соответствует следующий граф, изображенный на рисунке 4.1.*

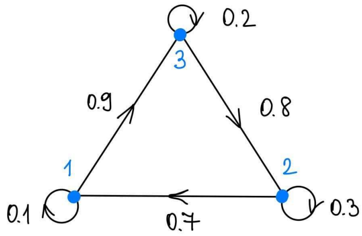

Рис. 4.1: Граф, описывающий цепь Маркова

Упражнение 4.7. *Для однородной цепи Маркова, порожденной начальным распределение μ и стохастической матрицей P найдите вероятность перехода за s шагов $P(X_{t+s} = j|X_t = i)$ и распределение вероятностей для состояния в момент времени t, т.е. $P(X_t = j)$.*

4.1.3 Эргодическая теорема

Вернемся к нашему примеру с отделами магазина. Предположим, покупатель первый раз пришел в магазин, он еще не привык к расположению отделов, поэтому перемещается по ним случайным образом. Однако с течением времени он изучит расположение отделов и привыкнет; можно посмотреть, как часто он посещает каждый из отделов, т.е. вычислить вероятности $\lim\limits_{s\to\infty} P(X_s = j)$. Давайте изучим вопрос о существовании этих пределов.

Определение 4.4. *Стохастическая матрица P называется эргодической, если существует такое T, что $(P^{(T)})_{ij} = p_{ij}^{(T)} > 0$ для всех i, j.*

Заметим, что поскольку множество возможных состояний $\Xi = 1, 2, ..r$ конечно, можно найти такое α что $(P^{(T)})_{ij} = p_{ij}^{(T)} \geq \alpha$ для всех i, j.

Упражнение 4.8. *Приведите пример неэргодической матрицы и соответствующей ей цепи Маркова.*

Теорема 4.2. *(Эргодическая теорема) Пусть цепь Маркова порождена начальным распределением μ и эргодической матрицей P. Тогда верны следующие утверждения:*

1. *переходные вероятности за s шагов сходятся к распределению π, т.е. $\lim\limits_{s\to\infty} p_{ij}^{(s)} = \pi_j$,*
2. *распределение вероятностей $P(X_s = j)$ для случайной величины X_s также сходится к π независимо от начального распределения μ, т.е. $\lim\limits_{s\to\infty} (\mu P^{(s)})_j = \pi_j$*
3. *такое распределение π единственно и инвариантно относительно стохастической матрицы P, т.е. $\pi P = \pi$.*

Определение 4.5. *Распределение π из предыдущей теоремы называется инвариантным или стационарным.*

Для доказательства теоремы нам понадобится следующая лемма.

Лемма 4.3. *Пусть $\mathcal{P} = \{\mu \in \mathbb{R}^{1\times r} : \sum\limits_{i=1}^{r} \mu_i = 1, \mu_i \geq 0 \quad \forall i\}$ - пространство вероятностей на Ξ. Тогда $d(\mu', \mu'') := \frac{1}{2}\sum\limits_{i=1}^{r} \mid \mu_i' - \mu_i'' \mid$ задает метрику на $\mathcal{P}$, при этом пространство $(\mathcal{P}, d)$ является полным метрическим пространством, т.е. выполняется критерий Коши. Также выполены следующие свойства*

$$d(\mu', \mu'') \leq 1, \tag{4.5}$$

$$d(\mu' Q, \mu'' Q) \leq d(\mu', \mu''), \tag{4.6}$$

где Q - стохастическая матрица.

$$d(\mu' Q, \mu'' Q) \leq (1-\alpha) d(\mu', \mu''), \tag{4.7}$$

где Q - стохастическая матрица, такая что $q_{ij} \geq \alpha \quad \forall i, j$.

Мы будем использовать эту лемму для доказательства теоремы, в частности применим формулу 4.7 для $Q = P^{(T)}$, где T такое, что $(P^{(T)})_{ij} = p_{ij}^{(T)} \geq \alpha$ для всех i, j (см. определение эргодической матрицы).

Доказательство. Пусть $\mu', \mu'' \in \mathcal{P}$. Обозначим через $\sum^+$ суммирование по тем индексам, для которых слагаемые положительны.

1. Для начала сформулируем несколько технических утверждений:

$$0 = \sum_{i=1}^{r} \mu_i' - \sum_{i=1}^{r} \mu_i'' = \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'') - \sum_{i=1}^{r}{}^{+} (\mu_i'' - \mu_i'),$$

$$d(\mu', \mu'') = \frac{1}{2} \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'') + \frac{1}{2} \sum_{i=1}^{r}{}^{+} (\mu_i'' - \mu_i') = \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'').$$

2. Ясно, что $d(\mu', \mu'') \leq 1$, т.к.

$$d(\mu', \mu'') := \frac{1}{2} \sum_{i=1}^{r} \mid \mu_i' - \mu_i'' \mid \leq \frac{1}{2} \sum_{i=1}^{r} \mid \mu_i' \mid + \frac{1}{2} \sum_{i=1}^{r} \mid \mu_i'' \mid \leq 1.$$

3. Пусть Q - стохастическая матрица. Тогда $\mu' Q, \mu'' Q$ будут распределениями вероятностей. Используя доказанные выше технические утверждения, запишем оценку на $d(\mu' Q, \mu'' Q)$:

$$d(\mu' Q, \mu'' Q) = \sum_{j=1}^{r}{}^{+} (\mu' Q - \mu'' Q)_j = \sum_{j \in J} (\mu' Q - \mu'' Q)_j =$$

$$= \sum_{j \in J} \sum_{i=1}^{r} (\mu_i' q_{ij} - \mu_i'' q_{ij}) = \sum_{i=1}^{r} (\mu_i' - \mu_i'') \sum_{j \in J} q_{ij} \leq \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'') \sum_{j \in J} q_{ij}.$$

В общем случае $\sum\limits_{j \in J} q_{ij} \leq \sum\limits_{j=1}^{r} q_{ij} = 1$. Тогда

$$d(\mu' Q, \mu'' Q) \leq \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'') = d(\mu', \mu''),$$

В случае, если $q_{ij} \geq \alpha \quad \forall i, j$, мы можем записать следующие неравенства:

$$\sum_{j\in J} q_{ij} = 1 - \sum_{j\notin J} q_{ij} \leq 1-\alpha,$$

$$d(\mu' Q, \mu'' Q) \leq (1-\alpha) \sum_{i=1}^{r}{}^{+} (\mu_i' - \mu_i'') = (1-\alpha) d(\mu', \mu'').$$

4. Утверждение, что $d(\mu', \mu'') := \frac{1}{2} \sum_{i=1}^{r} \mid \mu_i' - \mu_i'' \mid$ задает метрику, следует из того, что $d_1(m', m'') = |m' - m''|$ является метрикой в $\mathbb{R}$. Критерий Коши для d аналогичен критерию Коши для d_1.

$\square$

Доказательство. (Доказательство теоремы)

1. Для начала докажем, что распределение для X_s сходится при $s \to \infty$. Пусть μ - распределние X_0, тогда $\mu^s := \mu P^{(s)}$ - распределение для X_s. Применим формулу 4.7 для $Q = P^{(T)}$.

$$d(\mu^s, \mu^{s+t}) = d(\mu P^{(s)}, \mu P^{(s+t)}) \leq (1-\alpha) d(\mu P^{(s-T)}, \mu P^{(s+t-T)}).$$

   Повторим эту операцию m раз, где m такое, что $0 < n - mT < s$. Получим, что

$$d(\mu^s, \mu^{s+t}) \leq (1-\alpha)^m d(\mu P^{(s-mT)}, \mu P^{(s+t-mT)}) \leq (1-\alpha)^m.$$

   Для любого сколь угодно малого $\varepsilon > 0$ можно найти такие s, m, что $d(\mu^s, \mu^{s+t}) \leq (1-\alpha)^m < \varepsilon$. Таким образом последовательность $\{\mu_s\}$ является последовательностью Коши, а следовательно сходится, т.к. по лемме $(\mathcal{P}, d)$ - полное метрическое пространство.

2. Давайте проверим, что $\pi P = \pi$.

$$\pi P = \lim_{s\to\infty} \mu P^{(s)} P = \lim_{s\to\infty} \mu P^{(s+1)} = \pi.$$

3. Теперь давайте докажем, что такое π единственно.

   Пусть $\pi_1 P = \pi_1$, $\pi_2 P = \pi_2$. Тогда

$$d(\pi_1, \pi_2) = d(\pi_1 P^{(T)}, \pi_2 P^{(T)}) \leq (1-\alpha) d(\pi_1, \pi_2).$$

   Т.е. $d(\pi_1, \pi_2) = 0$, что означает, что π_1 и π_2 совпадают и распределения $\mu^s = \mu P^{(s)}$ сходятся к π независимо от начального распределения μ.

4. Теперь проверим, что переходные вероятности за s шагов сходятся к тому же пределу. Для этого возьмем $\mu = (0, 0, ...0, 1, 0, ...0)$, где единица стоит на i-ой позиции. Тогда переходные вероятности можно выразить следующим образом: $p_{ij}^{(s)} = (\mu P^{(s)})_j$. Нам уже известно, что $\mu_0 P^{(s)}$ сходится к π, т.е. $\lim\limits_{s\to\infty} (\mu P^{(s)})_j = \pi_j$. Тогда

$$\lim_{s\to\infty} p_{ij}^{(s)} = \lim_{s\to\infty} (\mu_0 P^{(s)})_j = \pi_j.$$

□

Пример 4.2. *Классическими примерами неэргодических цепей Маркова являются несвязные цепи маркова (число сообщающихся классов больше одного) и цепи с циклами.*

Теорема 4.4. *Условие детального баланса $\pi_i P_{ij} = \pi_j P_{ji}$ является достаточным условием стационарности π.*

Это теорема доказывается прямой проверкой уравнения $\pi P = \pi$. Заметим, что условие детального баланса играет ключевую роль в построении алгоритмов семплирования с помощью цепей Маркова, т.к. для того, чтобы получить семпл с целевым распределением π, нужно сконструировать цепь Маркова с переходными вероятностями P, удовлетворяющими условию баланса, а сходимость обеспечивается эргодической теоремой.

4.2 Обобщение цепи Маркова на случай непрерывного времени

Для практических приложений бывает полезно обобщить цепи Маркова на непрерывное время. Такие процессы могу описывать, например, процесс Пуассона (количество случайных событий, происходящее с заданной интенсивностью). Для того, чтобы определить такой процесс, нам снова понадобится воспользоваться теоремой Колмогорова, которая позволяет задать случайный процесс набором согласованных мер. Для начала рассмотрим, как выглядели согласованные меры в случае обычной цепи Маркова.

$$P\left(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_n = i_n\right) = \mu_{i_0} p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{n-1} i_n}.$$

$$F(t_1, .., t_m; A_{t_1}, .., A_{t_m}) = P\left(X_{t_1} \in A_{t_1}, .., X_{t_m} \in A_{t_m}\right) =$$

$$= \sum_{i_{t_1} \in A_{t_1}, ..., i_{t_m} \in A_{t_m}, i_0 \in \Xi, ... i_{t_{n-1}} \in \Xi} P\left(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_n = i_n\right)$$

$$= \sum_{i_{t_1} \in A_{t_1}, ..., i_{t_m} \in A_{t_m}, i_0 \in \Xi, ... i_{t_n} \in \Xi} \mu_{i_0} p_{i_0 i_1} p_{i_1 i_2} ... p_{i_{n-1} i_n}.$$

Заметим, что суммирование ведется по всем индексам, но i_{t_j}, которые параметризуют F, суммируются по A_{t_j}, а остальные - по всему допустимому пространству состояний. Формулы выглядят довольно громоздко, но оказывается, что для случая непрерывного времени они значительно упрощаются. Если мы решаем отказаться от дискретности времени, то вероятности перехода из состояния i_k в состояние i_{k+1} за 1 шаг следует заменить на такую же вероятность, но теперь за время $t_{k+1} - t_k$. В таком случае согласованные меры будут выглядеть следующим образом:

$$F(t_1, .., t_m; A_{t_1}, .., A_{t_m}) = \sum_{i_{t_1} \in A_{t_1}, ..., i_{t_m} \in A_{t_m}} \mu_{i_0} p_{i_0 i_{t_1}}(t_1 - 0) p_{i_{t_1} i_{t_2}}(t_2 - t_1) ... p_{i_{t_{n-1}} i_{t_m}}(t_m - t_{m-1}) \tag{4.8}$$

Заметим, что согласованные меры для цепи Маркова с непрерывным временем сильно напоминают согласованные меры для Виеровского процесса, отличие лишь в том, что вместо интегрирования мы использовали суммирование, т.к. пространство состояний остается дискретным.

$$\nu_{t_1,..,t_k}(A_1 \times .. \times A_k) := \int_{A_1 \times .. \times A_k} p(t_1, \tilde{x}, x_1) p(t_2 - t_1, x_1, x_2) ... p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 ... dx_k \ .$$

Осталось проверить, что введенные нами конечномерные распределения удовлетворяют условию согласованности, что по теореме Колмогорова является достаточным условием для существования случайного процесса с заданными конечномерными вероятностями. Делается это по аналогии с согласованностью мер для Винеровского процесса.

Теорема 4.5. *Меры, определенные формулой 4.8, удовлетворяют условию согласованности, если матрицы $P(t)$ являются стохастическими и образуют полугруппу, т.е. $P(s)P(t) = P(s + t)$.*

Полугруповвое свойство интуитивно означает следующее: комбинация перехода за s шагов и перехода за t шагов эквивалентно переходу за $s + t$ шагов. Явное доказательство согласованности оставляет читателю.

Упражнение 4.9. *Проверьте условие согласованности мер 4.8.*

Упражнение 4.10. *Проверьте непрерывность $P(t)$.*

Заметим, что конечность дискретного пространства состояний не вносила до сих пор каких-либо значительных ограничений, поэтому цепь Маркова можно определять с для счетного числа состояний. Осмысление этого факта оставляем читателю.

Определение 4.6. *Цепь Маркова в непрерывном времени с стохастическими матрицами перехода за t шагов $P(t)$, удовлетворяющими полугрупповому свойству, и начальным распределением μ - это случайный процесс $(X_t)_{t \in \mathbb{R}_+}$ со значениями в $\mathbb{Z}_+$, имеющий конечномерные распределения, удовлетворяющие формуле 4.8.*

Упражнение 4.11. *Проверьте, что цепь Маркова в непрерывном времени удовлетворяет следующему свойству:*

$$P(X_{t_n} = j | X_{t_0} = i_0, X_{t_1} = i_1, ...X_{t_{n-1}} = i) = p_{ij}(t_n - t_{n-1}), \tag{4.9}$$

где $p_{ij}(t)$ - это элементы матрицы $P(t)$.

По аналогии с обычной цепью Маркова можно задать стационарное распределение для цепи Маркова с непрерывным временем и сформулировать эргодическую теорему.

Определение 4.7. *Стационарное распределение для цепи Маркова с непрерывным временем, заданное переходными вероятностями $P(t)$ и начальным распределением μ, это распределение π, такое, что $\pi P(t) = \pi, \quad \forall t \geq 0$.*

Теорема 4.6. *Пусть $P(t) > 0$ для некоторого t. Тогда существует единственное стационарное распределение.*

Упражнение 4.12. *Докажите эргодическую теорему для цепи Маркова с непрерывным временем.*

4.2.1 Инфинитезимальная матрица

Рассмотрим цепи Маркова с дифференцируемоq в нуле переходной матрицей $P(t)$.

Определение 4.8. *Инфинитезимальная матрица Q для полугруппы $P(t)$ задается следующим образом $Q = \lim_{t \to 0+} \frac{P(t)-I}{t}$.*

Теорема 4.7. *Если Инфинитезимальная матрица Q существует, то существует $P'(t) = P(t)Q = QP(t)$. Данное дифференциальное уравнение вместе с начальным условием $P(t) = I$ имеет решение $P(t) = exp(tQ)$.*

Эта теорема доказывается прямым применением полугруппового свойства.

Теорема 4.8. *Если существует инфинитезимальная матрица, то уравнение $\pi Q = 0$ является необходимым и достаточным условием стационарности распределения π.*

Для доказательства достаточности обратитесь к определению матричной экспоненты, необходимости - к определению матрицы Q.

Пример 4.3. *Процесс Пуассона (марковская цепь с переходными вероятностями $p_{ij}(t) = Pois(j - i; \lambda t))$) обладает следующей инфинитезимальной матрицей:*

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & ... \\ 0 & -\lambda & \lambda & 0 & ... \\ 0 & 0 & -\lambda & \lambda & ... \\ ... & ... & ... & ... & ... \end{pmatrix}.$$

.