

# ФУНКЦИИ ПРИБЛИЖЕНИЯ ЗНАЧЕНИЯ ДЛЯ ЧАСТИЧНО НАБЛЮДАЕМЫХ МАРКОВСКИХ ПРОЦЕССОВ ПРИНЯТИЯ РЕШЕНИЯ

**Писарева Наталья Дмитриевна**

Аспирант, Курский государственный университет  
nata.pisareva.96@mail.ru

## VALUE APPROXIMATION FUNCTIONS FOR PARTIALLY OBSERVED MARKOV DECISION PROCESSES

**N. Pisareva**

*Summary.* This paper explores the problem of applying value approximation functions in the context of partially observable Markov decision processes (POMDPs, or POMDPs). POMDPs are a powerful tool for modeling situations in which decision making is based on imperfect information and the possible consequences of those decisions.

*Keywords:* Markov solutions, POMDP, process modeling.

*Аннотация.* В данной работе исследуется проблема применения функций приближения значения в контексте частично наблюдаемых Марковских процессов принятия решения (POMDP, или ЧНМППР). POMDP являются мощным инструментом для моделирования ситуаций, в которых принятие решений основано на несовершенной информации и возможных последствиях этих решений.

*Ключевые слова:* марковские решения, POMDP, моделирование процессов.

**А**ктуальность исследования частично наблюдаемых Марковских процессов принятия решения (POMDP) в настоящее время обусловлена рядом факторов.

Во-первых, современные проблемы и ситуации принятия решений все чаще обусловлены ограниченной доступностью информации и наличием неопределенности. Примеры таких ситуаций можно пронаблюдать в различных областях, например, на финансовых рынках, в сфере риск-менеджмента, смарт-городов, автономных систем и др. POMDP позволяют учесть эти факторы и разрабатывать оптимальные стратегии принятия решений, основываясь на имеющейся, но неполной информации.

Во-вторых, развитие компьютерных технологий и возможностей обработки больших объемов данных создало новые перспективы для применения POMDP. Алгоритмы и методы, основанные на POMDP, позволяют обрабатывать сложные и многомерные данные, а также учитывать динамические изменения состояний системы. Это имеет важное значение в контексте принятия решений в реальном времени и адаптивного управления.

В-третьих, POMDP являются активной областью исследований в математике, теории управления, искусственном интеллекте и других дисциплинах. Усовершенствование методов и подходов к приближению значений в POMDP позволяет создавать более точные и эффективные модели принятия решений. Это, в свою очередь, способствует развитию новых алгоритмов машинного обучения, адаптивных систем и смарт-технологий.

Одной из ключевых задач в рамках POMDP является оценка значения состояний и принятие оптимальных решений на основе имеющейся информации. Тем не менее, точный расчет значений может быть сложным или даже невозможным в плане вычислений из-за размерности состояний и ограничений на доступность информации. В этом смысле функции приближения значения играют важную роль, предлагая эффективные методы аппроксимации значений состояний в POMDP.

### Цель исследования

Целью данного исследования является изучение и анализ различных функций приближения значения для POMDP. Исследование направлено на разработку эффективных методов приближения, которые позволят принимать оптимальные решения в условиях неопределенности и ограниченной доступности информации.

### Проблема исследования

Одной из главных проблем, связанных с POMDP, является сложность определения оптимальных стратегий принятия решений из-за неопределенности и ограниченной информации. Традиционные методы, основанные на полностью наблюдаемых Марковских процессах принятия решений, неэффективны в случае частично наблюдаемых ситуаций. Поэтому необходимо исследовать и разработать новые функции приближения значения, способные учитывать особенности POMDP и преодолеть эти проблемы.

### Метод исследования

Исследование и анализ литературы с помощью баз данных, представленных в открытом доступе.

## Введение

Марковские процессы принятия решений (MDP, или МППР) — это математические модели, которые используются для моделирования ситуаций принятия последовательности решений в условиях неопределенности и стохастичности. Они представляют собой формализованный способ описания последовательности состояний системы и связанных с ними действий, а также оценки их влияния на будущие состояния.

В МППР основной принцип заключается в том, что решения принимаются последовательно в зависимости от текущего состояния системы и стохастических факторов, которые могут влиять на будущие состояния. Ключевой концепцией МППР является модель Маркова, согласно которой будущее состояние системы зависит только от текущего состояния и совершенного действия, а не от всей истории состояний и решений.

МППР находят применение во многих областях, включая экономику, финансы, управление производством, техническую диагностику, управление робототехнологиями и др. Они позволяют моделировать и оптимизировать процесс принятия решений в условиях неопределенности и учитывать различные факторы, включая вероятности перехода из одного состояния в другое, функции стоимости или награды за различные действия, а также ограничения и предпочтения принимающего решения.

Марковские процессы принятия решений условно делятся на полностью наблюдаемые (ПНМППР), где все состояния системы наблюдаются напрямую, и частично наблюдаемые (ЧНМППР), где доступна только неполная информация о состояниях. ЧНМППР представляют особый интерес, т. к. они позволяют моделировать реальные ситуации, где принятие решений основано на ограниченной информации, и требуют разработки специальных методов и алгоритмов для эффективного принятия оптимальных решений.

### Важность функций приближения значений к POMDP

Функции приближения значения для частично наблюдаемых Марковских процессов принятия решения (ЧНМППР) играют важную роль в эффективном принятии решений в условиях ограниченной доступности информации. Ниже перечислены основные сферы их применения:

1. Учет неопределенности: ЧНМППР позволяют учитывать неопределенность в будущих состояниях системы и стохастические факторы, которые могут влиять на принятие решений. Функции приближения значения позволяют оценить по-

тенциальные результаты различных действий и выбрать оптимальное решение, учитывая вероятности и возможные исходы.

2. Ограниченная доступность информации: В реальных ситуациях принятия решений не всегда доступна полная информация о текущем состоянии системы. ЧНМППР дают возможность моделировать такие ситуации, где доступна только неполная информация или наблюдения. Функции приближения значения здесь помогают оценить скрытые состояния системы на основе имеющейся информации.
3. Вычислительная эффективность: Точный расчет значений состояний в ЧНМППР может быть невозможным из-за издержек по доступу к информации. Функции приближения значения предлагают методы аппроксимации, которые позволяют оценить значения состояний системы при разумных вычислительных затратах.
4. Принятие решений в реальном времени: В некоторых ситуациях требуется принятие решений в режиме реального времени. Функции приближения значения позволяют быстро оценить оптимальные решения на основе доступной информации и актуального состояния системы, что особенно важно для систем с высокой динамикой и быстрыми изменениями состояний.

Использование функций приближения значения для ЧНМППР способствует более эффективному принятию решений в условиях ограниченной информации и неопределенности, позволяя учесть различные факторы и выбрать оптимальные стратегии действий. Они являются важным инструментом для моделирования и анализа принятия решений в различных областях, где необходимо учитывать неопределенность и принимать решения в режиме реального времени. Кроме того, функции приближения значения имеют важное значение в разработке алгоритмов машинного обучения и искусственного интеллекта.

### Процесс приближения значения для POMDP

Процесс приближения значения для частично наблюдаемых Марковских процессов принятия решения (ЧНМППР) может осуществляться различными методами и подходами, в зависимости от конкретной задачи и доступных ресурсов. Ниже перечислены несколько основных подходов, которые широко используются для этого:

1. Методы динамического программирования: Одним из наиболее распространенных подходов является применение методов динамического программирования, особенно методов итераций по состояниям или по времени. Эти методы позволяют рекурсивно оценивать значения состояний системы на основе имеющейся информации

и предыдущих оценок. Для ЧНМППР используются разные алгоритмы, в т.ч. алгоритм прямого прохода-обратного прохода и алгоритмы сглаживания, которые аппроксимируют и уточняют значения состояний.

2. Методы Монте-Карло: еще одним подходом является использование методов Монте-Карло, которые основаны на случайных выборках для оценки значений состояний. Эти методы генерируют большое количество случайных траекторий системы и на их основе аппроксимируют значения состояний и ожидаемых выигрышей. Примерами таких методов являются методы Монте-Карло с итерацией по политике или по значению.
3. Приближенные методы и алгоритмы: В случае сложных ЧНМППР, когда точное вычисление значений состояний затруднительно или невозможно, могут применяться приближенные методы и алгоритмы. Примерами таких подходов являются методы аппроксимации функций, включая аппроксимацию с использованием базисных функций (например, функций Гаусса или полиномов) и нейронных сетей, а также методы оптимизации, например, методы градиентного спуска или стохастического приближения.

Важно отметить, что для максимальной эффективности процесс приближения значения для ЧНМППР требует тщательного выбора метода и подхода с учетом характеристик системы, доступной информации и требуемой точности оценок.

### Заключение

В данной работе была рассмотрена проблема функций приближения значения для частично наблюдаемых

Марковских процессов принятия решения. Цель исследования заключалась в анализе и обсуждении важности и применения этих функций в контексте ограниченной доступности информации и неопределенности.

В процессе исследования было выявлено, что функции приближения значения играют важную роль в эффективном принятии решений в условиях частичной наблюдаемости. Они позволяют учесть неопределенность будущих состояний системы, ограниченную информацию и принимать решения в режиме реального времени. Они также являются неотъемлемой частью моделирования и анализа систем в различных областях, включая экономику, управление производством, финансы, техническую диагностику и др.

Было выяснено, что процесс приближения значения для ЧНМППР может успешно осуществляться на практике с использованием методов динамического программирования, методов Монте-Карло и приближенных методов. Выбор конкретного метода зависит от характеристик системы, требуемой точности оценок и доступных ресурсов.

Актуальность исследования подтверждается растущим интересом к моделированию и принятию решений в условиях ограниченной информации. Результаты исследования могут быть полезными для разработки новых алгоритмов и методов принятия решений, а также для применения в практических приложениях, где необходимо учитывать стохастические факторы и ограниченную доступность информации. Полученные результаты исследования могут служить основой для дальнейших исследований и разработки новых моделей и алгоритмов в области принятия решений.

### ЛИТЕРАТУРА

1. POMDP: Introduction to Partially Observable Markov Decision Processes; Michael Hahsler and Hossein Kamalzadeh; The R Project for Statistical Computing, 2021.
2. Approximation Methods for Partially Observed Markov Decision Processes (POMDPs); Caleb M. Bowyer; Cornell University, 2021; <https://doi.org/10.48550/arXiv.2108.13965>.
3. Approximations for Partially Observed Markov Decision Processes; Naci Saldi, Tamás Linder, Serdar Yüksel; Systems & Control: Foundations & Applications book series (SCFA), 2018.
4. A primer on partially observable Markov decision processes (POMDPs); Iadine Chadès, Luz V. Pascal, Sam Nicol, Cameron S. Fletcher, Jonathan Ferrer-Mestres; British Ecological Society, Volume 12, Issue 11, November 2021, Pages 2058–2072; <https://doi.org/10.1111/2041-210X.13692>.
5. Learning and Solving Partially Observable Markov Decision Processes; Guy Shani; Ben-Gurion University of the Negev Department of Computer Science; Dissertation submitted in partial fulfillment of the requirements for Ph.D. degree; 2007.

© Писарева Наталья Дмитриевна (nata.pisareva.96@mail.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»