

Р. С. Федорук (Великий Новгород, НовГУ). **Оценка параметров конечной цепи Маркова с дисконтируемыми доходами.**

Цепи Маркова с доходами являются удобным математическим объектом описания систем, соответствующих процессам окружающего мира. Зачастую перед некоторым исследователем может встать проблема нахождения параметров уже функционирующей системы с неизвестными параметрами с целью дальнейшего управления данной системой для получения наибольшего дохода. Интерес вызывает описание подхода к нахождению оптимального времени оценки параметров такой системы в условиях неизвестных, но постоянных параметров.

Рассматривается процесс с бесконечным временем планирования и дисконтируемыми доходами при значениях коэффициента дисконтирования β , близких к 1 ($0 < \beta < 1$). При значениях β , близких к 1, обесценивание доходов происходит со скоростью, достаточно малой для того, чтобы затраченное на определение параметров время T не привело к сильному обесцениванию доходов. Требуется найти оптимальное решение для каждого из состояний системы.

Управление в данной задаче условно можно разделить на 2 этапа. На первом этапе для определения вероятностных законов p работы системы производится подсчет количества переходов из состояния i в состояние j при выборе решения k за n шагов системы, обозначаемых $V_{ij}^k(n)$. На данном этапе управление состоит в том, что если система в момент времени n находится в состоянии i , то выполняется сравнение $V_i^k(n) = \sum_{j \in S} V_{ij}^k(n)$ при всех k и выбирается решение, которому соответствует минимальное значение $V_i^k(n)$, а если таких значений несколько, то любое из них (например, с наименьшим номером k или равновероятно). Важную роль на данном этапе играет оценка минимального значения времени T , которое следует затрачивать на проведение испытаний, достаточного для того, чтобы полученные по итогам испытаний оценки были нетривиальными и чтобы обеспечивалось выполнение неравенства $V_i^k(n) > 0$. На втором этапе в момент времени T , по формулам $\hat{p}_{ij}^k(n) = V_{ij}^k(n)/V_i^k(n)$, выполняются оценки переходных вероятностей, которые имеют смысл при $V_i^k(n) > 0$. На основании полученных таким образом параметров и алгоритма Ховарда определяется оптимальная стационарная стратегия, которая начиная с момента (шага) T применяется для оптимального управления системой.

В случае, если доходы системы являются детерминированными, то для их определения достаточно, чтобы система побывала в каждом из состояний, а если нет, то их оценку следует проводить по алгоритму, описанному для оценки вероятностных законов системы.

Моделирование на основании изложенного подхода показывает, что существует также зависимость между временем T и величиной коэффициента β и что при прочих равных условиях оптимальное время оценки параметров системы будет различным для различных коэффициентов дисконтирования.

СПИСОК ЛИТЕРАТУРЫ

1. Майн Х., Осаки С. Марковские процессы принятия решений. М.: Наука, 1977.
2. Колмогоров А. В., Федорук Р. С. Модель управления конечной цепью Маркова с дисконтируемыми доходами в случае неизвестных параметров. — Вестник Новгородского гос. ун-та. Сер. техн. науки, 2009, № 50, с. 29–30.