Численное решение динамической задачи стимулирования исполнителей в дискретном времени с учетом эффекта обучения

О.В. Павлов[1]

[1]Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

Аннотация. В работе рассматривается задача стимулирования исполнителей в дискретном времени с учетом эффекта обучения. Задача формулируется как динамическая игра между руководителем и исполнителями. Для решения задачи применен принцип компенсации затрат, который сводит исходную задачу к задаче оптимального управления в дискретном времени. Получены численные решения задачи для различных моделей кривых обучения с помощью метода динамического программирования Беллмана. Проведено исследование влияния ставки дисконтирования на решения динамической задачи стимулирования

1. Введение

В статье рассматривается игровая динамическая задача стимулирования исполнителей, выполняющих производственное задание в условиях освоения новой продукции. Для освоения новой продукции на промышленных предприятиях характерен эффект кривой обучения, который заключается в том, что затраты времени работников (трудоемкость) на выполнение многократно повторяющихся производственных операций снижаются.

Задача стимулирования исполнителей является одной из важнейших в теории управления организационными системами. Исполнитель (агент) выбирает действие (объем работы) исходя из своих экономических интересов. Руководство (центр) должно выбрать такую систему стимулирования на основе прогноза действий агента, что бы обеспечить выполнение своих экономических интересов.

Динамические игровые модели рассматриваются в теории активных систем [1], информационной теории иерархических систем [2–4] и теории динамических игр [5–7]. Задача динамического стимулирования агентов в терминах теории динамических игр называется обратная игра Штакельберга. Динамические обратные игры Штакельберга рассматриваются в научных публикациях [8–10]. В информационной теории иерархических систем задача динамического стимулирования получила название игра Гермейера Γ_2. Современный обзор полученных результатов информационной теории иерархических систем приводится в [4].

В теории активных систем [1] предлагается подход, основанный на принципе компенсации затрат агента. Центр компенсирует затраты агента в случае выбора оптимальной плановой траектории центра и не выплачивает материального вознаграждения в противном случае. Исходная задача разделяется на два задачи: выбор системы стимулирования и решение задачи оптимального управления. В работе [11] приводятся результаты, обобщающие теоремы из монографии [1].

В информационной теории иерархических систем [2–4] предлагается подход, связанный с выбором центром программы совместных действий с агентом и наказание за отклонение от этой программы. При этом центр исходит из принципа максимального гарантированного результата. В результате исходная задача стимулирования сводится к задаче оптимального управления.

Постановка и численное решение динамической задачи стимулирования агентов в условиях эффекта кривой обучения приведено автором в [12]. Данная статья является продолжением исследований и посвящена анализу влияния дисконтирования на решения динамической задачи стимулирования.

2. Постановка и алгоритм решения динамической задачи стимулирования исполнителей

Рассматривается двухуровневая динамическая производственная система, состоящая из центра и n независимых агентов. Агенты производят детали, из которых затем собирается готовое изделие. Трудовые затраты и материальное стимулирование агентов зависят только от их собственных действий. В работе применяется принцип декомпозиции игры [1], который позволяет рассматривать управление i-ым агентом независимо и не учитывать взаимодействие агентов между собой. Состояние динамической производственной системы зависит от действий агентов, а центр оказывает влияние на управляемую систему только через выплату материального вознаграждения агентам.

Динамика производства детали i-ым агентом описывается дискретным уравнением:

$$x_t = x_{t-1} + u_t, \quad t = \overline{1,T},$$

где x_t - кумулятивный объём производства детали за t-ый временной период, t – номер временного периода, u_t – объём производства детали в периоде t, T - количество рассматриваемых временных периодов.

До начала серийного производства известно количество произведенных деталей:

$$x_0 = X_0,$$

В конечный временной период кумулятивный объём деталей должен быть равен заданному:

$$x_T = X_0 + R,$$

где R – заданное количество деталей.

На объём производства детали наложены ограничения:

$$0 \le u_t \le X_0 + R - x_{t-1}, \; t = \overline{1,T}.$$

Целевой функцией центра является максимизация дисконтированной суммарной разности между доходом от произведенных деталей и затратами на материальное вознаграждение агента:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t}[pu_t - \sigma(x_t)] \to max, \tag{1}$$

где p – цена детали, $\sigma(x_t)$ - функция стимулирования центра, r – ставка дисконтирования.

Функция стимулирования центра – это правило, в соответствие с которым назначается материальное вознаграждение агенту за произведенный объем работы. Центр управляет производственным процессом через механизм материального стимулирования $\sigma(x_t)$, экономически побуждая агентов выполнять плановые объемы производства.

С помощью ставки дисконтирования учитываются временные предпочтения центра (агента) по стоимости денежных потоков. Чем более отдален по времени денежный поток, тем он для центра (агента) стоит дешевле.

Целевой функцией агента является максимизация дисконтированной суммарной разницы между материальным вознаграждением и трудовыми затратами, выраженными в денежной форме:

$$J_a = \sum_{t=1}^{T} \frac{1}{(1+r)^t} [\sigma(x_t) - C_t(u_t, x_{t-1})] \to max, \quad (2)$$

где r - ставка дисконтирования агента, $C_t(u_t, x_{t-1})$ – трудовые затраты агента.

Трудовые затраты агента определяются:

$$C_t(u_t, x_{t-1}) = sc_t u_t \quad (3)$$

где s – стоимость одного норма часа для агента, c_t - трудоемкость изготовления детали.

Зависимость трудоемкости изготовления детали от кумулятивного объёма производства описывается различными моделями кривой обучения, приведенными в научной литературе [13]-[16].

В соответствии со своими экономическими интересами агент выбирает объёмы производства деталей, которые максимизируют его целевую функцию (2).

Задача управления центра заключается в выборе оптимальной системы стимулирования $\sigma(x_t)$, при которой агент будет изготавливать такие объёмы производства деталей u_t^{opt}, которые максимизируют целевую функцию центра (1).

Для решения сформулированной задачи управления применяется принцип компенсации затрат [1,11]. Алгоритм решения состоит в разделении исходной задачи на две задачи: выбор компенсаторной системы стимулирования и решение задачи оптимального управления с целевой функцией равной разности между доходом центра и трудовыми затратами агента.

1.Выбор компенсаторной системы стимулирования.

Центр выбирает компенсаторную систему стимулирования, которая заключается в компенсации затрат агента в случае выбора оптимального планового объема производства центра $x_t^{opt}, t = \overline{1,T}$ и отсутствия материальных выплат в противном случае:

$$\sigma(x_t) = \begin{cases} C_t(u_t, x_{t-1}), \; если \; x_t = x_t^{opt}, \; для \; \forall \; t = \overline{1,T}, \\ 0, \qquad если \; x_t \neq x_t^{opt}, \; для \; \forall \; t = \overline{1,T}. \end{cases}$$

2.Решение задачи оптимального управления с целевой функцией равной разности между доходом центра и трудовыми затратами агента.

Что бы побудить агента выбрать плановый объем производства центр выплачивает материальное вознаграждение равное затратам агента:

$$\sigma(x_t) = C_t(u_t, x_{t-1}). \quad (4)$$

Подставим в целевую функцию центра формулу (4) с учетом (3):

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} [p - sc_t] u_t \to max.$$

Так как цена детали p постоянная, то центр может увеличить свою прибыль, только за счет минимизации суммарных затрат на выплату материального вознаграждения агента. Целевая функция центра примет вид:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} sc_t u_t \to min.$$

Таким образом, исходная игровая задача динамического стимулирования сведена к задаче оптимального управления:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} sc_t u_t \to min. \quad (5)$$

$$x_t = x_{t-1} + u_t, \quad t = \overline{1,T}, \quad (6)$$

$$x_0 = X_0, \quad (7)$$

$$x_T = X_0 + R, \quad (8)$$

$$0 \le u_t \le X_0 + R - x_{t-1}, \; t = \overline{1,T}. \quad (9)$$

Задача центра заключается в выборе оптимальных объёмов производства деталей $u_t^{opt}, t = \overline{1,n}$, с учетом ограничения (9), при которых производственный процесс (6) перейдет из начального состояния (7) в конечное состояние (8) и будет достигнут минимум целевой функции центра (5).

Сформулированная задача оптимального управления (5)-(9) была решена с помощью метода динамического программирования Беллмана [17]-[18], реализованного на языке программирования pascal.

3. Результаты численного решения динамической задачи стимулирования агента

Численное решение задачи оптимального управления проведено на примере производства деталей предприятия АО «Салют». По данным предприятия построены регрессионные модели трудоемкости изготовления деталей: степенная, экспоненциальная и логистическая.

Степенная модель трудоемкости:

$$c_t = 42{,}64 x_{t-1}^{-0{,}3}.$$

Экспоненциальная модель трудоемкости:

$$c_t = 9{,}17 + 6{,}16 e^{-0{,}03 x_{t-1}}.$$

Логистическая модель трудоемкости:

$$c_t = 55{,}10 + 36{,}61 \left[\frac{1}{1 + 0{,}017 e^{0{,}05 x_{t-1}}} \right].$$

Для решения задачи использовались следующие данные: количество временных периодов T=12 месяцев, объём производства деталей R=240 шт., производственный опыт до серийного производства $x_0 = 1$ шт. Дискретный шаг изменения объема производимых агентом деталей в методе динамического программирования – 1 шт.

Численные решения задачи оптимального управления для степенной, экспоненциальной и логистической моделей трудоемкости представлены на рисунках 1-3. На рисунках изображены оптимальные траектории кумулятивных объемов производства для различных ставок дисконтирования.

Из анализа рисунков 1-2 следует, что для степенной и экспоненциальной моделей трудоемкости оптимальной траекторией кумулятивного объема производства является выпуклая кривая. Оптимальной стратегией центра является перераспределение больших объемов производства деталей на последние временные периоды, в которых трудоемкость производства деталей меньше, чем в начальных.

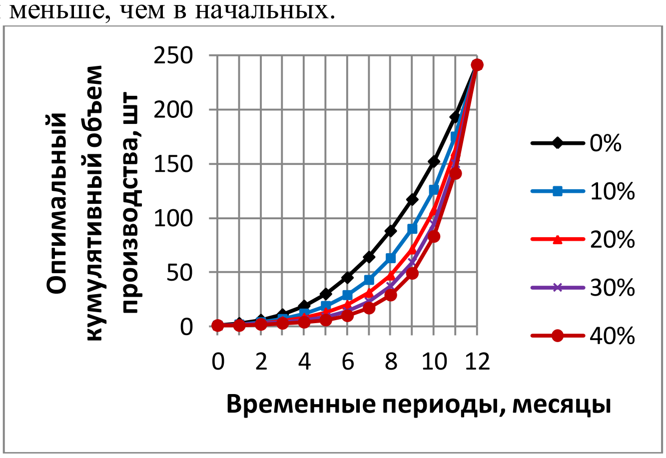

Рисунок 1. Зависимость оптимального кумулятивного объёма производства от ставки дисконтирования для степенной модели трудоемкости.

С увеличением ставки дисконтирования стратегия центра по перераспределению больших объемов производства деталей на последние временные периоды усиливается. Это объясняется более «дешевой» стоимостью денег, которые центр выплачивает агенту в качестве материального вознаграждения в отдаленных временных периодах. При больших значениях ставки дисконтирования возникает эффект откладывания производства деталей с начальных временных периодов на более поздние. Центру экономически выгодно отложить производство деталей на поздние временные периоды, так как в этом случае его суммарные дисконтированные затраты будут минимальны.

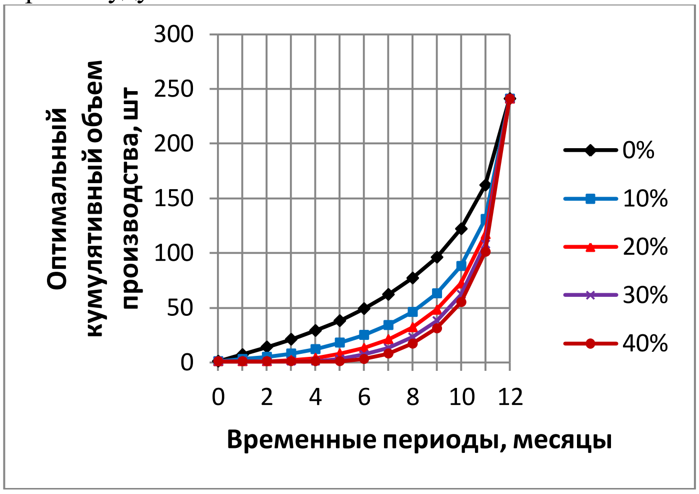

Рисунок 2. Зависимость оптимального кумулятивного объёма производства от ставки дисконтирования для экспоненциальной модели трудоемкости.

Анализируя рисунок 3, приходим к выводу, что для логистической модели трудоемкости при отсутствии дисконтирования (r=0%) оптимальной траекторией кумулятивного объема производства является логистическая кривая. Оптимальная траектория кумулятивного объема производства состоит из двух участков: вогнутого и выпуклого.

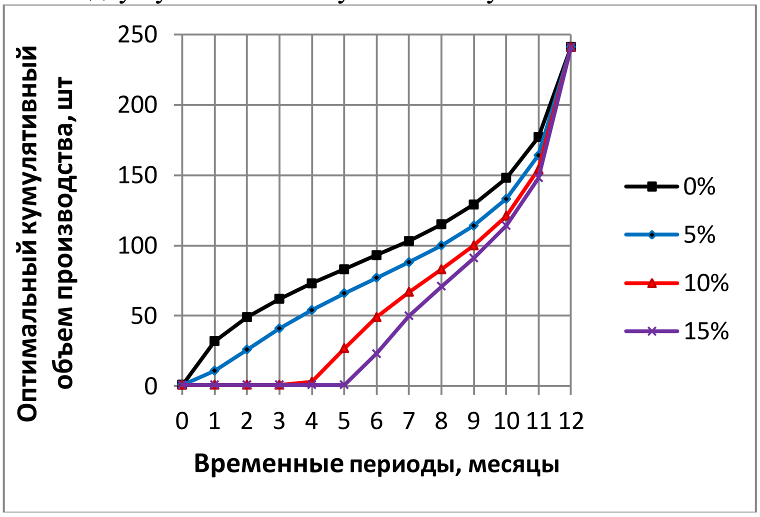

Рисунок 3. Зависимость оптимального кумулятивного объёма производства от ставки дисконтирования для логистической модели трудоемкости.

На рисунке 4 приведены оптимальные траектории объемов производства для различных ставок дисконтирования r. Оптимальной стратегией центра при отсутствии дисконтирования (r=0%) является: уменьшение объемов производства для вогнутого участка оптимальной траектории кумулятивного объема производства и увеличение объемов производства для выпуклого участка траектории. Минимум объема производства соответствует точке перегиба оптимальной траектории кумулятивного объема производства.

При учете дисконтирования для логистической модели трудоемкости также наблюдается эффект откладывания производства деталей с начальных временных периодов на более поздние. Учет дисконтирования приводит к появлению на оптимальной траектории кумулятивного объема производства дополнительного выпуклого участка в начальных временных периодах.

Оптимальная траектория кумулятивного объема производства трансформируется в кривую из трех участков: выпуклого, вогнутого и выпуклого. Оптимальная стратегия центра состоит: на выпуклых участках траектории в увеличении объемов производства, на вогнутых - в уменьшении. Точкам перегиба соответствуют экстремальные значения объемов производства.

Рисунок 4. Зависимость оптимального объёма производства от ставки дисконтирования для логистической модели трудоемкости.

4. Заключение

В работе рассмотрена динамическая задача стимулирования исполнителей в дискретном времени с учетом эффекта кривой обучения.

Для решения задачи применен принцип компенсации затрат, который состоит в разделении исходной задачи на две задачи: выбор компенсаторной системы стимулирования и решение задачи оптимального управления с целевой функцией равной разности между доходом центра и трудовыми затратами агента.

С помощью метода динамического программирования Беллмана получены численные решения задачи оптимального управления для различных моделей трудоемкостей. Проведено исследование влияния ставки дисконтирования на решения задачи.

На основе численного исследования сформулированы следующие выводы:

1.Оптимальной стратегией центра для степенной и экспоненциальной моделей кривых обучения является перераспределение больших объемов производства деталей на последние временные периоды, в которых трудоемкость производства деталей меньше, чем в начальных.

2.Учет дисконтирования для степенной и экспоненциальной моделей кривых обучения приводит к еще большему перераспределению объемов производства деталей на последние временные периоды.

3. Учет дисконтирования для всех рассмотренных моделей кривых обучения приводит к появлению эффекта откладывания производства с начальных периодов на более поздние.

4. Оптимальной траекторией кумулятивного объема производства в случае логистической модели кривой обучения является кривая, состоящая из нескольких выпуклых и вогнутых участков. Оптимальная стратегия центра состоит: в увеличении объемов производства на выпуклых участках траектории, на вогнутых - в уменьшении. Точкам перегиба соответствуют экстремальные значения объемов производства.

5.Учет дисконтирования для логистической модели кривой обучения приводит к перераспределению объемов производства деталей на средние и последние временные периоды.

5. Благодарности
Исследование выполнено при финансовой поддержке РФФИ и Самарской области в рамках научного проекта № 17-46-630606.

6. Литература

[1] Новиков, Д.А. Механизмы управления динамическими активными системами / Д.А. Новиков, И.М. Смирнов, Т.Е. Шохина – М.: ИПУ РАН, 2002. – 124 с.
[2] Горелик, В.А. Анализ конфликтных ситуаций в системах управления / В.А. Горелик, М.А. Горелов, А.Ф. Кононенко – М.: Радио и связь, 1991. – 288 с.
[3] Горелик, В.А. Теоретико-игровые модели принятия решений в эколого-экономических системах / В.А. Горелик, А.Ф. Кононенко – М.: Радио и связь, 1982. – 144 с.
[4] Горелов, М.А. Динамические игры. III. Иерархические игры / М.А. Горелов, А.Ф. Кононенко // Автоматика и телемеханика. – 2015. – Т. 2. – С. 89-106.
[5] Basar, T. Dynamic Noncooperative Game Theory / T. Basar, G.J. Olsder – Philadelphia: SIAM, 1999. – 519 p.
[6] Dockner, E. Differential games in economics and management Science / E. Dockner, S. Jorgensen, N.V. Long, G. Sorger – Cambridge: Cambridge University Press, 2000. – 382 p.
[7] Угольницкий, Г.А. Управление устойчивым развитием активных систем – Ростов-на-Дону: Издательство Южного федерального университета, 2016. – 940 с.
[8] Olsder, G.J. Phenomena in inverse Stackelberg games. Part 2: dynamic problems // Journal of Optimization Theory and Applications. – 2009. – Vol. 143(3). – P. 601–618.
[9] Groot, N. Reverse Stackelberg games. Part I: basic framework / N. Groot, B. De Schutter, H. Hellendoorn // Proceedings of the IEEE International Conference on Control Applications, 2012. – P. 421-426.
[10] Groot, N. Reverse Stackelberg games. Part II: results and open issues / N. Groot, B. De Schutter, H. Hellendoorn // Proceedings of the IEEE International Conference on Control Applications, 2012. – P. 427-432.
[11] Рохлин, Д.Б. Равновесие Штакельберга в динамической модели стимулирования с полной информацией / Д.Б. Рохлин, Г.А. Угольницкий // Автоматика и телемеханика. – 2018 – № 4. – С. 152-166.
[12] Павлов, О.В. Численное исследование проблемы согласованного управления проектами по освоению новой продукции // Экономические науки. – 2018. Т. 167, № 10. – С. 41-49.
[13] Wright, T.P. Factors affecting the cost of airplanes // Journal of the aeronautical sciences. – 1936. – Vol. 3(4). – P. 122-128.
[14] Yelle, L.E. The learning curve: Historical review and comprehensive survey / L.E. Yelle // Decision Sciences. – 1979. – Vol. 10(2). – P. 302-328.
[15] Badiru, A. Computational survey of univariate and multivariate learning curvemodels // IEEE Transactions on Engineering Management. – 1992. – Vol. 39(2). – P. 176-188.
[16] Jaber, M.Y. Learning Curves: Theory, Models, and Applications – Boca Raton: CRC Press, 2011. – 476 P.
[17] Беллман, Р. Динамическое программирование – М.: Издательство иностранной литературы, 1960.
[18] Калихман, И.Л. Динамическое программирование в примерах и задачах / И.Л. Калихман, М.А. Войтенко – М.: Высш. школа, 1979.

Numerical solution of the dynamic incentive problem in discrete time with account the learning curve effect

O.V. Pavlov[1]

[1]Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

Abstract. The paper considers the of the dynamic incentive problem in discrete time with account the learning curve effect. The task is formulated as a dynamic game between the leader and the performers. To solve the problem, the principle of cost recovery is applied, which reduces the original task to the optimal control problem in discrete time. Numerical solutions of the problem for various models of learning curves are obtained using the Bellman dynamic programming method. The study of the impact of the discount rate on the solution of the incentive problem was conducted.