Часть I МАТРИЧНЫЕ ИГРЫ

Рассмотрим игру, в которой участвуют два игрока, причем каждый из игроков имеет конечное число стратегий. Обозначим для удобства одного из игроков через А, в другого — через *B*.

Предположим, что игрок *A* имеет *m* стратегий — A_1, A_2, ..., A_m, а игрок *B* имеет *n* стратегий B_1, B_2, ..., B_n.

Пусть игрок *A* выбрал стратегию A_i, а игрок *B* стратегию B_k. Будем считать, что выбор игроками стратегий A_i и B_k однозначно определяет исход игры — выигрыш a_{ik} игрока *A* и выигрыш b_{ik} игрока *B*, причем эти выигрыши связаны равенством

$$b_{ik} = -a_{ik}$$

(отрицательный выигрыш на бытовом языке обычно называют проигрышем).

Последнее условие показывает, что в рассматриваемых обстоятельствах выигрыш одного из игроков равен выигрышу другого, взятому с противоположным знаком. Поэтому при анализе такой игры можно рассматривать выигрыши только одного из игроков. Пусть это будут, например, выигрыши игрока *A*.

Если нам известны значения a_{ik} выигрыша при каждой паре стратегий (в каждой ситуации) { A_i, B_k }, i = 1, 2,..., m, k = 1, 2,..., n, то их удобно записывать или в виде прямоугольной таблицы, строки которой соответствуют стратегиям игрока *A*, а столбцы – стратегиям игрока *B*:

	B_1	B_1	…	B_n
A_1	a_{11}	a_{11}		a_{11}
A_2	a_{11}	a_{11}		a_{11}
A_m	a_{m1}	a_{11}		a_{mn}

или в виде матрицы

Полученная матрица имеет размер m x n и называется матрицей игры, или платежной матрицей (отсюда и название игры — матричная).

Рассматриваемую игру часто называют игрой m x n или m x n игрой.

Замечание. Матричные игры относятся к разряду так называемых антагонистических игр, то есть игр, в которых интересы игроков прямо противоположны.

Пример 1. Каждый из двух игроков A и B одновременно и независимо друг от друга записывает на листе бумаги любое целое число. Если выписанные числа имеют одинаковую четность, то игрок A получает от игрока B 1 рубль, а если разную, то наоборот – игрок A платит 1 рубль игроку B.

У игрока A две стратегии: A_1 – записать четное число и A_2 – записать нечетное число. У игрока B такие же две стратегии: B_1 – записать четное число и B_2 – записать нечетное число. Выбор игроками соответственно стратегий A_i и B_k однозначно определяет исход игры – выигрыш игрока A.

Матрица этой 2x2 игры имеет следующий вид

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

(здесь строки соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B).

1. Равновесная ситуация

Рассмотрим следующий пример.

Пример 2. Два игрока A и B, не глядя друг на друга, кладут на стол по картонному кружку красного (r), зеленого (g) или синего (b) цветов, сравнивают цвета кружков и расплачиваются друг с другом так, как покапано в матрице игры

$$\begin{pmatrix} -2 & 2 & -1 \\ 2 & 1 & 1 \\ 3 & -3 & 1 \end{pmatrix}$$

(напомним, что у этой 3 x 3-матрицы строки соответствуют стратегиям игрока A, а столбцы - стратегиям игрока B).

Считая, что эта 3 x 3 игра повторяется многократно, попробуем определить оптимальные стратегии каждого из игроков.

Начнем с последовательного анализа стратегий игрока A, не забывая о том, что, выбирая стратегию игрока A, должно принимать в расчет, что его противник B может ответить на нее той из своих стратегий, при которой выигрыш игрока A будет минимальным.

Так, на стратегию A, он ответит стратегией B_r, (минимальный выигрыш равен -2, что на самом деле означает проигрыш игрока A, равный 2), на стратегию A_g – сстратегией B_g или B_b (минимальный выигрыш игрока A равен 1), а на стратегию A_b – стратегией B_g (минимальный выигрыш игрока A равен -3).

Запишем эти минимальные выигрыши в правом столбце таблицы:

	B_r	B_g	B_b	
A_r	-2	2	-1	-2
A_g	2	1	1	1
A_b	3	-3	1	-3

maxmin. Неудивительно, что игрок A останавливает свой выбор на стратегии A_g, при которой его минимальный выигрыш максимален (из трех чисел -2, 1 и -3 максимальным является 1):

	B_r	B_g	B_b	
A_r	-2	2	-1	-2

A_g	2	1	1	**1**
A_b	3	-3	1	-3

maxmin = 1.

Если игрок A будет придерживаться этой стратегии, то ему гарантирован выигрыш, не меньший 1, при любом поведении противника в игре.

Аналогичные рассуждения можно провести и за игрока B. Так как игрок B заинтересован в том, чтобы обратить выигрыш игрока A в минимум, то ему нужно проанализировать каждую свою стратегию с точки зрения максимального выигрыша игрока A.

Выбирая свою стратегию, игрок B должен учитывать, что при этом стратегией его противника A может оказаться та, при которой выигрыш игрока A будет максимальным. Так, на стратегию B_r он ответит стратегией A_b (максимальный выигрыш игрока A равен 3), на стратегию B_g – стратегией A_r (максимальный выигрыш игрока A равен 2), а на стратегию B_b – стратегией A_g или A_b (максимальный выигрыш игрока A равен 1). Эти максимальные выигрыши записаны в нижней строке таблицы

	B_r	B_g	B_b	
A_r	-2	2	-1	-2
A_g	2	1	1	1
A_b	3	-3	1	-3
	3	2	1	

minmax. Неудивительно, вели игрок B остановит свой выбор на стратегии B_b, при которой максимальный выигрыш игрока A минимален (из трех чисел 3, 2 и 1 минимальным является 1):

	B_r	B_g	B_b	
A_r	-2	2	-1	-2
A_g	2	1	1	1
A_b	3	-3	1	-3
	3	2	**1**	

minmax = 1.

Если игрок B будет придерживаться этой стратегии, то при любом поведении противника он проиграет не больше 1.

В рассматриваемой игре числа maxmin и minmax совпали:

maxmin = minmax = 1

(соответствующие элементы в таблице

	B_r	B_g	$\boldsymbol{B_b}$
A_r	-2	2	-1
$\boldsymbol{A_g}$	2	1	1
A_b	3	-3	1

выделены жирным шрифтом).

Выделенные стратегии A_g и B_b являются оптимальными стратегиями игроков A и B,

$$\boxed{A_g = A_{opt},\ B_b = B_{opt}}$$

в следующем смысле:

при многократном повторении игры отказ от выбранной стратегии любым из игроков уменьшает его шансы на выигрыш (увеличивает шансы на проигрыш).

В самом деле, если игрок A будет придерживаться не стратегии A_{opt}, а выберет иную стратегию, например, A_r, то вряд ли стоит рассчитывать на то, что игрок B этого не заметит. Конечно, заметит и не преминет воспользоваться своим наблюдением. Ясно, что в этом случае он отдаст предпочтение стратегии B_r. А на выбор A_b игрок B ответит, например, так – B_g. В результате отказа от стратегии A_g, выигрыш игрока A уменьшится.

Если же от стратегии B_{opt} отказывается игрок B, выбирая, например, стратегию B_r, то игрок A может ответить на это стратегией A_b и, тем самым, увеличить свой выигрыш. В случае стратегии B_g ответ игрока A – A_r .

Тем самым, ситуация { A_g, B_b } оказывается равновесной.

Еще раз подчеркнем, что элементами матрицы игры являются числа, описывающие выигрыш игрока A. Более точно, выигрыш соответствует положительному элементу платежной матрицы, а отрицательный указывает на проигрыш игрока A.

Матрица выплат игроку B получается из матрицы игры заменой каждого ее элемента на противоположный по знаку.

Рассмотрим теперь произвольную матричную игру

$$A = \begin{pmatrix} a_{11} & a_{12} & \vdots & a_{1n} \\ a_{21} & a_{22} & \vdots & a_{2n} \\ .... & .... & . & .... \\ a_{m1} & a_{m2} & \vdots & a_{mn} \end{pmatrix}$$

(строки заданной m x n-матрицы соответствуют стратегиям игрока A, а столбцы - стратегиям игрока B) и опишем общий алгоритм, посредством которого можно определить, есть ли в этой игре ситуация равновесия или ее нет.

В теории игр предполагается, что оба игрока действуют разумно, то есть стремятся к получению максимального выигрыша, считая, что соперник действует наилучшим (для себя) образом.

Действия игрока А

1-й шаг. В каждой строке матрицы А ищется минимальный элемент

$$\alpha_i = \min_k a_{ik}\ ,\ i = 1, 2, \dots, m.$$

Полученные числа

$$\alpha_1, \alpha_2, \dots, \alpha_m$$

приписываются к заданной таблице в виде правого добавочного столбца

$$\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & \alpha_1 \\ a_{21} & a_{22} & \dots & a_{2n} & \alpha_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & \alpha_m \end{array}$$

Пояснение. Выбирая стратегию A_i, игрок A вправе рассчитывать на то, что в результате разумных действий противника (игрока B) он выиграет не меньше чем a_i.

2-й шаг. Среди чисел

$$\alpha_1, \alpha_2, \dots, \alpha_m$$

выбирается максимальное число

$$\alpha = \max_i \alpha_i$$

или, подробнее,

$$\alpha = \max_i \min_k a_{ik}$$

Специально отметим, что выбранное число α является одним из элементов заданной матрицы A.

Пояснение. Действуя наиболее осторожно и рассчитывая на наиболее разумное поведение противника, игрок A должен остановиться на той стратегии A_i, для которой число a_i, является максимальным.

Если игрок A будет придерживаться стратегии, Выбранной описанным выше способом, то при любом поведении игрока B игроку A гарантирован выигрыш, не меньший α.

Число α называется *нижней ценой игры*.

Принцип построения стратегии игрока A, основанный на максимизации минимальных выигрышей, называется *принципом максимина*, а выбираемая в соответствии с этим принципом стратегия A_i^0 — *максиминной стратегией* игрока A.

Действия игрока В

1 шаг. В каждом столбце матрицы A ищется максимальный элемент

$\beta_k = \max\limits_i a_{ik}$, $k = 1, 2, \ldots , n$.

Полученные числа

$$\beta_1, \beta_2, \ldots , \beta_n$$

приписываются к заданной таблице в виде нижней добавочной строки

$$\begin{array}{cccc|c}
a_{11} & a_{12} & \ldots & a_{1n} & \alpha_1 \\
a_{11} & a_{22} & \ldots & a_{2n} & \alpha_2 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
a_{11} & a_{m2} & \ldots & a_{mn} & \alpha_m \\
\hline
\beta_1 & \beta_2 & \ldots & \beta_n &
\end{array}$$

Пояснение. Выбирая стратегию B_k, игрок B должен рассчитывать на то, что в результате разумных действий противника (игрока A) он проиграет не больше чем β_k.

2-й шаг. Среди чисел

$$\beta_1, \beta_2, \ldots , \beta_n$$

выбирается минимальное число

$$\beta = \min_k \beta_k$$

или, подробнее,

$$\beta = \min_k \max_i a_{ik}$$

Выбранное число β также является одним из элементов заданной матрицы А.

Пояснение. Действуя наиболее осторожно и рассчитывая на наиболее разумное поведение противника, игрок B должен остановиться на той стратегии B_k*, для которой число β_k является минимальным.

Если игрок B будет придерживаться стратегии, выбранной описанным выше способом, то при любом поведении игрока A игроку B гарантирован проигрыш, не боль ший β.

Число β называется верхней ценой игры.

Принцип построения стратегии игрока B, основанный на минимизации максимальных потерь, называется *принципом минимакса*, а выбираемая в соответствии с этим принципом стратегия B_k^0 – минимаксной стратегией игрока B.

Нижняя цена игры α и верхняя цена игры β всегда связаны неравенством

$$\alpha \le \beta$$

Замечание. Реализация описанного алгоритма требует $2mn - 1$ сравнений элементов матрицы A:

$$(n - 1)m + m - 1 = mn - 1$$

сравнений для определения α,

$$(n - 1)m + m - 1 = mn - 1$$

сравнений для определения β и одно сравнение полученных чисел α и β.

Если

$$\alpha = \beta,$$

или, подробнее,

$$\max_i \min_k a_{ik} = a_{i^0 k^0} = \min_k \max_i a_{ik},$$

то ситуация $\{A_i^0, B_k^0\}$ оказывается равновесной, и ни один из игроков не заинтересован в том, чтобы ее нарушить (в этом нетрудно убедиться путем рассуждений, подобных проведенным при анализе игры в примере 2).

В том случае, когда нижняя цена игры равна верхней цене игры, их общее значение называется просто ценой игры и обозначается через v.

Цена игры совпадает с элементом $a_{i^0 k^0}$ матрицы игры A, расположенным на пересечении i^0-й строки (стратегия A_{i^0} игрока A) и k^0-го столбца (стратегия B_{k^0} игрока B) – минимальный в своей строке и максимальный в своем столбце.

Этот элемент называют *седловой точкой матрицы A*, или *точкой равновесия*, а про игру говорят, что она *имеет седловую точку*.

Стратегии A_{i^0} и B_{k^0}, соответствующие седловой точке, называются оптимальными, а совокупность оптимальных ситуаций и цена игры – *решением матричной игры с седловой точкой*.

Замечание. Седловых точек в матричной игре может быть несколько, но все они имеют одно и то же значение.

Матричные игры с седловой точкой важны и интересны, однако более типичным является случай, когда применение описанного алгоритма приводит к неравенству

$$\alpha < \beta.$$

Как показывает следующий пример, в этом случае предложенный выбор стратегий уже, вообще говоря, к равновесной ситуации не приводит, и при многократном ее повторении у игроков вполне могут возникнуть мотивы к нарушению рекомендаций, основанных на описанном алгоритме действий игроков A и B.

Пример 3. Рассмотрим 3 х 3 игру, заданную матрицей

$$A = \begin{pmatrix} 4 & -1 & -3 \\ -2 & 1 & 3 \\ 0 & 2 & -3 \end{pmatrix}.$$

Применив предложенный алгоритм

4	1	-3	-3
-2	1	3	-2
0	**2**	-3	-3
4	2	3	

находим нижнюю цену игры α = -2 и соответствующую ей стратегию A_2, верхнюю цену игры $\beta = 2$ и соответствующую ей стратегию B_2.

Нетрудно убедиться в том, что пока игроки придерживаются этих стратегий, средний выигрыш при многократном повторений игры равен 1. Он больше нижней мены игры, но меньше верхней цены.

Однако если игроку B станет известно, что игрок A придерживается стратегии A_2, он немедленно ответит стратегией B_1 и сведет его выигрыш к проигрышу -2. В свою очередь, на стратегию B_1 у игрока A имеется ответная стратегия A_1, дающая ему выигрыш 4.

Тем самым, ситуация $\{ A_2, B_2\}$ равновесной не является.

2. Смешанные стратегии

В случае, когда нижняя цена игры α и верхняя цена игры β не совпадают,

$$\alpha < \beta,$$

игрок A может обеспечить себе выигрыш, не меньший α, а игрок B имеет возможность не дать ему больше, чем β. Возникает вопрос – а как разделить между игроками разность

$$\beta - \alpha?$$

Предыдущие построения на этот вопрос ответа не дают — тесны рамки возможных действий игроков. Поэтому довольно ясно, что механизм, обеспечивающий получение каждым из игроков как можно большей доли этой разности, следует искать в определенном расширении стратегических возможностей, имеющихся у игроков изначально.

Оказывается, что компромиссного распределения разности $\beta - \alpha$ между игроками и уверенного получения каждым игроком своей доли при многократном повторении игры можно достичь путем случайного применения ими своих первоначальных, чистых стратегий. При таких действиях

- во-первых, обеспечивается наибольшая скрытность выбора стратегии (результат выбора не может стать известным противнику, поскольку он неизвестен самомуигроку),
- во-вторых, при разумном построении механизма случайного выбора стратегий, последние оказываются оптимальными.

Случайная величина, значениями которой являются стратегии игрока, называется его *смешанной стратегией*.

Тем самым, задание смешанной стратегии игрока состоит в указании тех вероятностей, с которыми выбираются его первоначальные стратегии.

Основные определения

Рассмотрим произвольную m x n игру, заданную m x n-матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} & \vdots & a_{1n} \\ a_{21} & a_{22} & \vdots & a_{2n} \\ \cdots & \cdots & . & \cdots \\ a_{m1} & a_{m2} & \vdots & a_{mn} \end{pmatrix}$$

Так как игрок A имеет m чистых стратегий, то его смешанная стратегия может, быть описана набором m неотрицательных чисел

$$p_1 \geq 0, p_2 \geq 0, \dots, p_m \geq 0,$$

сумма которых равна 1,

$$\sum_{i=1}^{m} p_i = 1.$$

Смешанная стратегия второго игрока B, имеющего n чистых стратегий, описывается набором n неотрицательных чисел

$$q_1 \geq 0, q_2 \geq 0, \dots, q_n \geq 0,$$

сумма которых равна 1,

$$\sum_{k=1}^{n} q_k = 1.$$

Замечание. Каждая чистая стратегия является частным случаем смешанной стратегии: в частности, чистая стратегия A_i является смешанной стратегией, описываемой набором чисел Рев котором $p_1, p_2, \dots, p_m$, в котором

$$p_i = 1, \quad p_j = 0 \quad (j \neq i)$$

Подчеркнем, что для соблюдения секретности каждый из игроков применяет свои стратегии независимо от другого игрока.

Таким образом, задав два набора

$$P = \{p_1, p_2, \dots, p_m\}, \qquad Q = \{q_1, q_2, \dots, q_n\},$$

мы оказываемся в ситуации *в смешанных стратегиях*. В этих условиях каждая обычная ситуация (в чистых стратегиях) $\{A_i, B_k\}$ по определению является случайным событием и, ввиду независимости наборов P и Q, реализуется с вероятностью $p_i q_k$. В этой ситуации $\{A_i, B_k\}$ игрок A получает выигрыш a_{ik}. Тем самым, математическое ожидание выигрыша игрока A в условиях ситуации в смешанных стратегиях (P, Q) равно

$$H_A(P,Q) = \sum_{i=1}^{m} \sum_{k=1}^{n} a_{ik} p_i q_k .$$

Это число и принимается за *средний выигрыш* игрока A при смешанных стратегиях

$$P = \{p_1, p_2, \dots, p_m\}, \qquad Q = \{q_1, q_2, \dots, q_n\},$$

Определение. Стратегии

$$P^0 = \{p_1^0, p_2^0, \dots, p_m^0\} \quad \text{и} \quad Q^0 = \{q_1^0, q_2^0, \dots, q_n^0\},$$

называются *оптимальными смешанными стратегиями* игроков A и B соответственно, если выполнено следующее соотношение

$$H_A(P,Q^0) \leq H_A(P^0,Q^0) \leq H_A(P^0,Q).$$

Пояснение. Выписанные неравенства означают следующее:

левое неравенство – отклонение игрока A от оптимальной стратегии P^0 при условии, что игрок B придерживается стратегии Q^0, приводит к тому, что выигрыш отклонившегося игрока A может только уменьшиться,

правое неравенство – отклонение игрока B от оптимальной стратегии Q^0 при условии, что игрок A придерживается стратегии P^0, приводит к тому, что выигрыш игрока A может только возрасти, и значит, выигрыш игрока B – только уменьшиться.

Приведенное условие оптимальности равносильно тому, что[1]

$$\max_P \min_Q H_A(P,Q) = H_A(P^0,Q^0) = \min_Q \max_P H_A(P,Q)$$

Величина

$$\upsilon = H_A(P^0,Q^0),$$

определяемая последней формулой, называется *ценой игры*.

Набор (P^0, Q^0, v), состоящий из оптимальных смешанных стратегий игроков A и B и цены игры, называется *решением матричной игры*.

[1] Экстремальные величины $\max_P \min_Q H_A(P,Q)$ $\min_Q \max_P H_A(P,Q)$ всегда существуют вследствие того, что функция $H_A(P,Q)$ является непрерывной на замкнутом множестве

$$p_i \geq 0, \quad \sum p_i = 1, \quad q_k \geq 0, \quad \sum q_k = 1.$$

Пример 4. Рассмотрим 2 х 2 матричную игру из примера 1. Матрица этой игры имеет следующий вид

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Как нетрудно убедиться, седловой точки у нее нет.

Смешанные стратегии игроков A и B могут быть описаны парами чисел

$$P = \{p,\ 1 - p\} \quad \text{и} \quad Q = \{q,\ 1 - q\}$$

соответственно.

Средний выигрыш игрока A вычисляется так

$$H_A(P,Q) = 1\cdot p\cdot q + (-1)\cdot p\cdot(1-q) + (-1)p\cdot(1-p)\cdot q + 2\cdot(1-p)(1-q)?$$

откуда легко следует, что

$$H_A(P,Q) = 4pq - 2p - 2q + 1 = (2p-1)(2q-1)$$

Последнее удобно записать так

$$H_A(P,Q) = 4\left(p-\frac{1}{2}\right)\left(q-\frac{1}{2}\right).$$

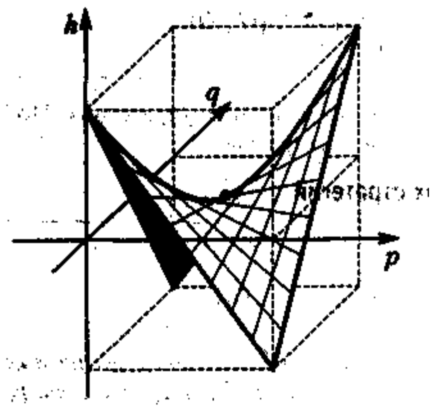


Рис. 1

Полученная формула показывает, что если игрок A в половине случаев записывает на листе бумаге четное (нечетное) число (выбирает p = 1/2), то независимо от того, что делает игрок B, ожидаемый (средний) выигрыш игрока A в каждой партии будет нулевым.

Если же игрок A выберет p > 1/2 (так что разность p-1/2 будет положительной), то узнав об этом, игрок B может выбрать q < 1/2 (так что разность q - 1/2 будет отрицательной) и, тем самым, сделать средний выигрыш игрока A отрицательным, то есть заставит его проиграть. Если же игрок A выберет p < 1/2 (так что разность p -1/2 будет отрицательной) и игрок B узнает об этом, то он может выбрать q > 1/2 (так что разность q -1/2 будет положительной) и вновь сделать выигрыш игрока A отрицательным, то есть опять заставит его проиграть.

Исследуем теперь эту формулу с точки зрения игрока B.

Если игрок A выбирает p = 1/2, то ожидаемый (средний) выигрыш игрока B независимо от его действий будет нулевым в каждой партии. Но если игрок B выберет q > 1/2 (так что разность q - 1/2 будет положительной), то, узнав об этом, игрок A может выбрать p < 1/2 (так что разность p - 1/2 будет отрицательной), и тогда игрок B будет проигрывать в каждой партии. Если же игрок B выберет q < 1/2 (так что разность q -1/2 будет отрицательной) и игрок A узнает об этом, то он может выбрать p > 1/2 (так что разность p -1/2 будет положительной) и вновь заставит игрока B проиграть.

Тем самым наборы

$$P^0 = \left\{\frac{1}{2}, \frac{1}{2}\right\},\ Q^0 = \left\{\frac{1}{2}, \frac{1}{2}\right\}$$

является оптимальными, а исход игры ничейным:

$$v = 0.$$

Замечание. На рисунке 1 показано, как устроена поверхность, описываемая функцией

$$h(p,q) = 4pq - 2p - 2q + 1 = (2p-1)(2q-1)$$

Точка

$$\left(\frac{1}{2}, \frac{1}{2}, 0\right)$$

является седловой точкой (точкой перевала) этой поверхности. Именно эта точка и дает решение рассматриваемой матричной игры.

Естественно возникают два *ключевых вопроса*:

1-й – какие матричные игры имеют решение в смешанных стратегиях?

2-й – как находить решение матричной игры, если оно существует?

Ответы на эти вопросы дают следующие две теоремы.

Основная теорема матричных игр

Теорема 1 (Дж. фон Нейман). Для матричной игры с любой матрицей A величины

$$\max_P \min_Q H_A(P,Q) \quad \min_Q \max_P H_A(P,Q)$$

равны между собой,

$$\max_P \min_Q H_A(P,Q) = \min_Q \max_P H_A(P,Q)$$

Более того, существует хотя бы одна ситуация в смешанных стратегиях (P^0, Q^0), для которой выполняется соотношение

$$H_A(P^0, Q^0) = \max_P \min_Q H_A(P,Q) = \min_Q \max_P H_A(P,Q)$$

Иными словами, любая матричная игра имеет решение в смешанных стратегиях. Поиск этого решения опирается на следующие установленные факты.

Основные свойства оптимальных смешанных стратегий

Теорема 2. Пусть

$$P^0 = \{p_1^0, p_2^0, ..., p_m^0\} \text{ и } Q^0 = \{q_1^0, q_2^0, ..., q_n^0\}$$

- оптимальные смешанные стратегии и v – цена игры,

Оптимальная смешанная стратегия P^0 игрока A смешивается только из тех чистых стратегий A_i, $i = 1, 2,..., m$, то есть отличными от нуля могут быть вероятности p_i только с теми номерами $i = 1,2, ... , m$, для которых выполнены равенства

$$\sum_{k=1}^{n} a_{ik} q_k^0 = v.$$

Это означает, что смешиваются не все чистые стратегии. Аналогично, в оптимальной смешанной стратегии Q^0 игрока B отличных от нуля могут быть только те вероятности q_k, для номеров $k = 1,2, ... , n$, которых выполнены равенства

$$\sum_{i=1}^{m} a_{ik} p_i^0 = v.$$

Кроме того, имеют место соотношения

$$v = \min_{1\le k\le n} \sum_{i=1}^{m} a_{ik} p_i^0 = \max_P \min_{1\le k\le n} \sum_{i=1}^{m} a_{ik} p_i = \min_Q \max_{1\le i\le m} \sum_{k=1}^{n} a_{ik} q_k = \max_{1\le i\le m} \sum_{k=1}^{n} a_{ik} q_k^0 = v.$$

В этом последнем скоплении равенств, по существу, и лежат истоки, питающие методы построения решений матричных игр. Опишем простейшие из них.

3. Методы решения матричных игр

Наши рассмотрения мы начнем с матричных игр, число стратегий хотя бы одного из игроков в которых равно двум.

Для построения решений 2 x n и m x 2 игр существует эффективный метод, основанный на простых геометрических соображениях. Этот метод, называют *графическим*.

2 x n игры

Пусть

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix}$$

- платежная матрица 2 x n игры.

Согласно теореме о двойном описании игры (теорема 2) нахождение цены игры и оптимального значения p^0 для игрока A равносильно разрешению уравнения

$$v = \min_{1\le k\le n}\left(a_{1k}p^0 + a_{2k}\left(1-p^0\right)\right) = \max_{0\le p\le 1}\min_{1\le k\le n}\left(a_{1k}p + a_{2k}(1-p)\right).$$

Опишем общую схему, приводящую к искомому результату.

Максимум функции

$$\min_{1\le k\le n}\left(a_{1k}p + a_{2k}(1-p)\right) \qquad (*)$$

проще всего найти, построив ее график. Для этого поступают следующим образом.

Предположим, что игрок A выбрал смешанную стратегию $P = \{p,\ 1 - p\}$, а игрок B - k-ю чистую стратегию, $k = 1, 2, \dots, n$. Тогда средний выигрыш игрока A в ситуации $\{P,\ k\}$ оказывается равным

$$(k):\quad \omega = \left(a_{1k}p + a_{2k}(1-p)\right).$$

На плоскости (p, ω) уравнение (k) описывает прямую. Тем самым, каждой чистой стратегия игрока B на этой плоскости соответствует своя прямая.

Поэтому сначала на плоскости (ω, p) последовательно и аккуратно рисуются все прямые

$$(k):\quad \omega = \left(a_{1k}p + a_{2k}(1-p)\right),\ k = 1, 2, \dots, n$$

(рис.2). Затем для каждого значения p, $0 \le p \le 1$, путем визуального сравнения соответствующих ему значений ω на каждой из построенных прямых определяется и отмечается минимальное из них (рис. 3). В результате описанной процедуры получается ломаная, которая и является графиком функции (*) (выделена жирным на рис. 4). Эта ломаная как бы огибает снизу все семейство построенных прямых,

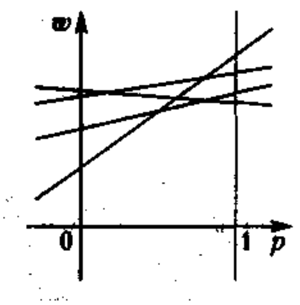


Рис.2

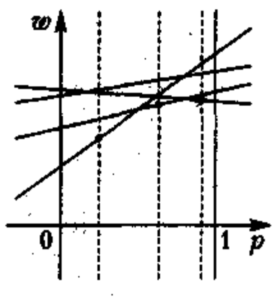


Рис.3

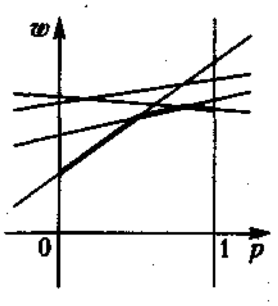


Рис.4

и по этой причине ее принято называть *нижней огибающей* этого семейства. Верхняя точка построенной нижней огибающей определяет и цену игры – v и оптимальную стратегию игрока A – $P^0 = \{p^0,\ 1 - p^0\}$ (рис. 5).

Замечание. Описанная процедура может рассматриваться как некоторый аналог максиминного подхода при отсутствии седловой точки.

Опробуем описанную схему решения 2 х n игры на конкретном примере.

Пример 5. Рассмотрим игру, заданную 2 х 6 матрицей

$$\begin{pmatrix} 6 & 4 & 3 & 1 & -1 & 0 \\ -2 & -1 & 1 & 0 & 5 & 4 \end{pmatrix}$$

Решение.

1-й шаг. *Анализ игры на наличие седловой точки.*

Нижняя цена игры равна -1, верхняя цена игры равна 1. Седловой точки нет. Решение игры нужно искать в смешанных стратегиях.

2-й шаг. *Вычисление средних выигрышей игрока А* (проводится при условии, что игрок *B* выбирает только чистые стратегии).

Из таблицы

p	6	4	3	1	-1	0
$1 - p$	-2	-1	1	0	5	4

легко получаем:

$$\begin{aligned} &(1): \omega = 6p - 2(1 - p), \\ &(2): \omega = 4p - (1 - p), \\ &(3): \omega = 3p + (1 - p), \\ &(4): \omega = p, \\ &(5): \omega = 6p - 2(1 - p), \\ &(6): \omega = 4(1 - p), \end{aligned}$$

3-й шаг. *Построение нижней огибающей.*

Аккуратно строим на координатной плоскости (p, ω) все шесть прямых, уравнения которых получены на 2-м шаге (рис. 6), и находим их нижнюю огибающую.

4-й шаг. *Отыскание цены игры и оптимальной смешанной стратегии игрока А.*

При аккуратном построении нижней огибающей, нетрудно определить, точкой пересечения каких двух из построенных шести прямых является ее наивысшая точка. В данном случае это прямые (4) и (5), заданные уравнениями $\omega = p$ и $\omega = -p + 5(1 - p)$ соответственно. Решая систему уравнений

$$\begin{aligned} &\omega = p, \\ &\omega = -p + 5(1 - p), , \end{aligned}$$

получаем

$$p^0 = \frac{5}{7}, \ \omega^0 = \frac{5}{7}$$

(рис.7).

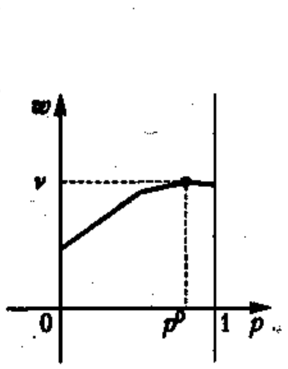


Рис.5

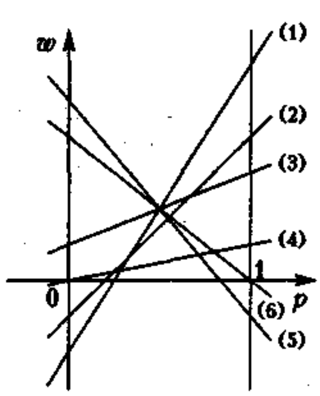


Рис.6

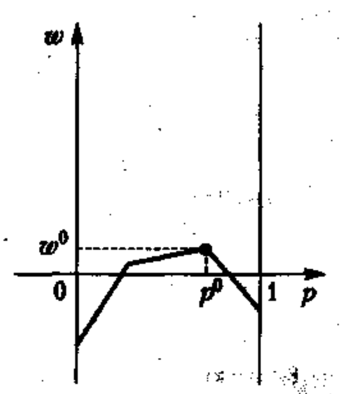


Рис.7

Тем самым, цена игры v и оптимальная стратегия P^0 игрока A соответственно равны

$$v = \frac{5}{7}, \; P^0 = \left\{\frac{5}{7}, \frac{2}{7}\right\}.$$

Собственно, этим и заканчивается решение игры для игрока A, поскольку его в первую очередь интересует отыскание собственной оптимальной стратегии и ожидаемого наилучшего гарантированного результата.

Замечание. Решающий матричную игру обычно отождествляет себя с одним из игроков (как правило, это игрок A), считая другого своим противником. Это связано еще и с тем, что в некоторых случаях основное внимание уделяется поиску оптимальных стратегий только игрока A, а стратегии противника могут вообще не интересовать исследователя.

Однако в целом ряде случаев оказывается важным знать оптимальные смешанные стратегии обоих игроков.

Как ищется оптимальная смешанная стратегия игрока A, мы уже описали. Покажем теперь, как отыскать оптимальную смешанную стратегию игрока B.

Здесь, в зависимости от формы нижней огибающей, может представиться несколько случаев.

А. Нижняя огибающая имеет ровно одну наивысшую точку (p^0, ω^0):

1) Если $p^0 = 0$ (оптимальная стратегия игрока А – чистая стратегия A_2), то игроку B выгодно применять чистую стратегию, соответствующую номеру прямой, проходящей через точку (0, ω^0) и имеющей наибольший отрицательный наклон (рис. 8).

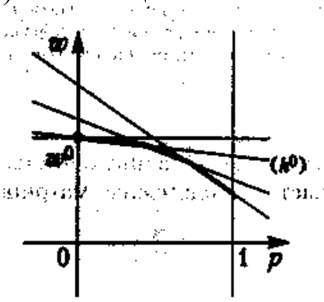


Рис.8

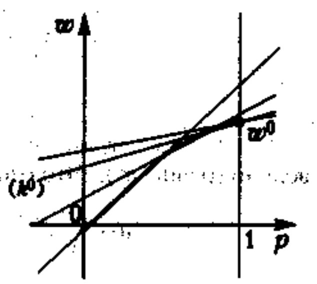


Рис.9

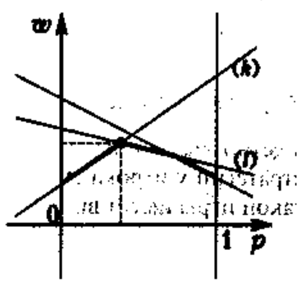


Рис.10

2) Если $p^0 == 1$ (оптимальная стратегия игрока A — чистая стратегия A_1), то оптимальной для игрока B является чистая стратегия, соответствующая номеру прямой, проходящей через точку (1, ω^0) и имеющей наименьший положительный наклон (рис. 9).

3) Если $0 < p^0 < 1$, то в наивысшей точке нижней огибающей пересекаются, по меньшей мере, две прямые, одна из которых (k-я) имеет положительный наклон, а другая (l-я) — отрицательный (рис. 10), и оптимальная смешанная стратегия игрока B получается, если положить

$$q_k = q,\; q_l = 1 - q,\; q_j = 0,\; j \neq k, l.$$

где q – решение уравнения

$$a_{1k}q + a_{1l}(1-q) = a_{2k}q + a_{2l}(1-q).$$

Б. Нижняя огибающая содержит горизонтальный участок, соответствующий чистой стратегии k^0 игрока B, которая и является оптимальной для него (рис. 11).

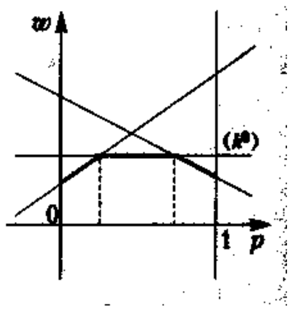


Рис. 11

Пример 5 (продолжение). Покажем теперь, как найти полное решение игры, то есть еще и оптимальную смешанную стратегию

$$Q^0 = \{q_1^0, q_2^0, q_3^0, q_4^0, q_5^0, q_6^0\}$$

игрока B.

Для этого поступают так:

1) полагают

$$q_1^0 = 0,\ q_2^0 = 0,\ q_3^0 = 0,\ q_4^0 = q,\ q_5^0 = 1 - q,\ q_6^0 = 0$$

(выделяя тем самым из шести чистых стратегий игрока В стратегии B_4 и B_5. которым соответствуют прямые (4) и (5), определяющие наивысшую точку нижней огибающей),

2) приравнивают любой из двух средних выигрышей игрока B (игрок A выбирает только чистые стратегии), отвечающих предложенной смешанной стратегии

0	0	0	q	1 - q	0
6	4	3	1	-1	0
-2	-1	1	0	5	4

к цене игры

$$q - (1 - q) = \frac{5}{7},\ 5(1 - q) = \frac{5}{7}$$

3) получают (в обоих случаях), что

$$q^0 = \frac{6}{7}.$$

Полное решение игры имеет следующий вид

$$P^0 = \left\{\frac{5}{7}, \frac{2}{7}\right\},\ Q^0 = \left\{0,\ 0,\ 0,\ \frac{6}{7},\ \frac{1}{7},\ 0\right\},\ \nu = \frac{5}{7}.$$

Замечание. Ситуация с наличием лишь двух конкурирующих стратегий игрока A нельзя считать надуманной. Она возникает сравнительно часто. Например, в случае, если нужно сравнить два образца некоторого изделия (скажем, старого и модернизированного) с целью выяснения возможности замены, это весьма удобно сделать при помощи платежной матрицы 2 х n.

m x 2 игры

Пусть теперь в матричной игре две чистые стратегии имеет игрок B, а число чистых стратегий у игрока A произвольно (равно m). Это означает, что платежная матрица такой игры имеет вид

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix}.$$

Анализ такой игры во многом напоминает рассуждения, описанные для игры 2 х n.

Пусть $Q = \{q, 1 - q\}$ — произвольная смешанная стратегия игрока B. Если игрок A выбирает i-ю чистую стратегию, $i = 1, 2, \ldots, n$, то средний выигрыш игрока B в ситуации $\{i, Q\}$ будет равным

$$(i)\colon\ \omega = a_{i1}q + a_{i2}(1-q)\quad i = 1, 2, \ldots, m. \qquad (*)$$

Зависимость этого выигрыша от переменной q описывается прямой.

Графиком функции

$$\max_{1 \le i \le m}\left(a_{i1}q + a_{i2}(1-q)\right)$$

является верхняя огибающая семейства прямых (*), соответствующих чистым стратегиям игрока A (рис. 12).

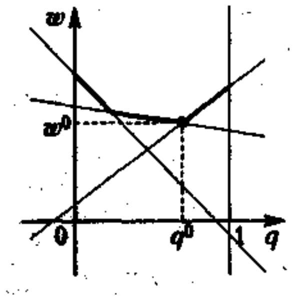


Рис.12

Абсциссой нижней точки полученной ломаной будет значение q^0, определяющее оптимальную смешанную стратегию игрока B, а ординатой ω^0 – цена игры.

Замечание. Отыскание оптимальной смешанной стратегии игрока А проводится по той же схеме, которая позволяет находить оптимальную смешанную стратегию игрока B в игре 2 х n.

Рассмотрим конкретный пример.

Пример 6. 3 х 2 игра задана матрицей

$$\begin{pmatrix} 3 & -1 \\ -1 & 3 \\ 1 & 0 \end{pmatrix}.$$

Решение.

1 шаг. *Анализ игры на наличие седловой точки*.

Нижняя цена игры равна 0, верхняя цена игры равна 3. Седловой точки нет. Решение игры нужно искать в смешанных стратегиях.

2-й шаг. *Вычисление средних выигрышей игрока B* (проводится при условии, что игрок A выбирает только чистые стратегии).

Из таблицы

q	$1 - q$
3	-1
-1	3
1	0

получаем:

(1): $\omega = 3q - (1 - q)$

(2): $\omega = -q + 3(1 - q)$

(3): $\omega = q$

3-й шаг. Построение верхней огибающей.

Построим на координатной плоскости (q, ω) все три прямых, а затем и их верхнюю огибающую (рис. 13).

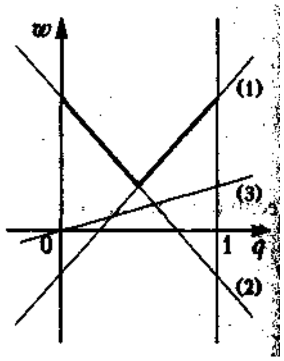


Рис. 13

4-й шаг. *Отыскание иены игры и оптимальной смешанной стратегии игрока B*.

Нижняя точка верхней огибающей является точкой пересечения прямых (1) и (2). Решая систему уравнений

$$\omega = 3q - (1 - q),$$
$$\omega = -q + 3(1 - q),$$

получаем

$$q^0 = \frac{1}{2}, \ \omega^0 = 1.$$

5-й шаг. *Отыскание оптимальной смешанной стратегии игрока A*.

Полагая

$$p_1^0 = p, \ p_2^0 = 1 - p, \ p_3^0 = 0,$$

приравниваем средние выигрыши игрока A, соответствующие чистым выигрышам игрока B,

$$3p - (1 - p) = -p + 3(1 - p),$$

и находим p^0 = 1/2.

Таким образом, цена игры и оптимальные смешанные стратегии игроков A и B соответственно

$$\nu = 1, \quad P^0 = \left\{\frac{1}{2}, \ \frac{1}{2}, \ 0\right\}, \quad Q^0 = \left\{\frac{1}{2}, \ \frac{1}{2}\right\}.$$

m x n игры

В принципе решение любой матричной юры сводится к решению стандартной задачи линейного программирования и, тем самым, может быть найдено методами линейного программирования. При этом требуемый объем вычислений напрямую зависит от числа чистых стратегий игроков (растет с его увеличением и, значит, с увеличением размеров матрицы игры). Поэтому любые приемы предварительного анализа игры, позволяющие уменьшать размеры ее платежной матрицы или еще как-то упрощать эту матрицу, не нанося ущерба решению, играют на практике весьма важную роль

Правило доминирования

В целом ряде случаев анализ платежной матрицы обнаруживает, что некоторые чистые стратегам не могут внести никакого вклада в искомые оптимальные смешанные стратегии. Отбрасывание подобных стратегий позволяет заменить первоначальную матрицу на матрицу выигрышей меньших размеров.

Опишем одну из таких возможностей более подробно.

$$A = \begin{pmatrix} a_{11} & a_{12} & \vdots & a_{1n} \\ a_{21} & a_{22} & \vdots & a_{2n} \\ \cdots & \cdots & \vdots & \cdots \\ a_{m1} & a_{m2} & \vdots & a_{mn} \end{pmatrix}$$

— произвольная m x n-матрица. Будем говорить, что i-я строка матрицы A

$$a_{i1} \ a_{i2} \ \ldots \ a_{in}$$

не больше j -и строки этой матрицы

$$a_{j1} \ a_{j2} \ \ldots \ a_{jn},$$

если одновременно выполнены следующие n неравенств

$$a_{i1} \le a_{j1}, \ \ a_{i2} \le a_{j2}, \ \ldots \ , \ \ a_{in} \le a_{jn}.$$

При этом говорят также, что j -я строка доминирует i-ю строку, или что стратегия A_j игрока A доминирует стратегию A_i.

Замечание. Игрок A поступит разумно, если будет избегать стратегий, которым в матрице игры отвечают доминируемые строки.

Если в матрице A одна из строк (j -я) доминирует другую строку (i-ю), то число строк в матрице A можно уменьшить путем отбрасывания доминируемой строки (i-й). Далее, будем говорить, что k-й столбец матрицы A

$$\begin{matrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{matrix}$$

не меньше l-го столбца этой матрицы

$$\begin{matrix} a_{1l} \\ a_{2l} \\ \vdots \\ a_{ml} \end{matrix}$$

если одновременно выполнены следующие m неравенств

$$a_{1k} \le a_{1l}, \ \ a_{2k} \le a_{2l}, \ \ldots \ , \ \ a_{mk} \le a_{ml}.$$

При этом говорят также, что 2-й столбец *доминирует* k-й столбец, или что стратегия B_l игрока *B доминирует* стратегию B_k.

Замечание. Игрок B поступит разумно, если будет избегать стратегий, которым в матрице игры отвечают доминируемые столбцы.

Если в матрице A один из столбцов (l-й) доминирует другой столбец (k-й), то число столбцов в матрице A можно уменьшить путем отбрасывания доминируемого столбца (k-го).

Важное замечание. Оптимальные смешанные стратегии в игре с матрицей, полученной усечением исходной за счет доминируемых строк и столбцов, дадут оптимальное решение в исходной игре: доминируемые чистые стратегии игроков в смешении не участвуют – соответствующие им вероятности следует взять равными нулю.

Пример 7. Рассмотрим игру с матрицей

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ -2 & 0 & 1 & 0 \\ 2 & 1 & -1 & -2 \\ -1 & 0 & 2 & 1 \end{pmatrix}.$$

Сравнивая строки матрицы, видим, что 1-я строка совпадает с 4-й строкой, или, что то же, стратегия A_4 дублирует стратегию A_1. Тем самым, одну из этих, строк можно вычеркнуть, не нанося ущерба решению

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ -2 & 0 & 1 & 0 \\ 2 & 1 & -1 & -2 \end{pmatrix}.$$

Поэлементна сравнивая 1-ю и 2-ю строки, замечаем, что 1-я строка доминирует 2-ю строку, или, что то же, стратегия A_1 доминирует стратегию A_2. Это вновь позволяет уменьшить число строк матрицы

$$\begin{pmatrix} -1 & 0 & 2 & 1 \\ 2 & 1 & -1 & -2 \end{pmatrix}$$

Замечая, что 4-й столбец полученной матрицы доминирует ее 3-й столбец, приходим к игре с 2 х 3-матрицей

$$\begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix}$$

Решая эту 2 х 3 игру графическим методом, находим ее решение – цену игры и оптимальные смешанные стратегии игроков A и B:

$$\nu = 0, \quad P^0 = \left\{\frac{2}{3}, \frac{1}{3}\right\}, \quad Q^0 = \left\{\frac{1}{2}, 0, \frac{1}{2}\right\}.$$

Возвращаясь к исходной 4 х 4 игре, получаем окончательный ответ:

$$\nu = 0, \quad \left\{\frac{2}{3}, \ 0, \ \frac{1}{3}, 0\right\}, \quad \left\{\frac{1}{2}, 0, 0, \frac{1}{2}\right\}$$

Замечание. При отбрасывании доминируемых строк и столбцов некоторые из оптимальных стратегий могут быть потеряны. Однако цена игры не изменится, и по усеченной матрице может быть найдена хотя бы одна пара оптимальных смешанных, стратегий.

Аффинное правило

При поиске решения матричных игр часто оказывается полезным следующее свойство.

Допустимые преобразования матрицы игры и ее цена. Оптимальные стратегии у матричных игр, элементы матриц A и C которых связаны равенствами

$$c_{ik} = \lambda a_{ik} + \mu, \quad i = 1, 2, \ldots, m; \quad k = 1, 2, \ldots, n,$$

где $\lambda > 0$, а μ – произвольно, имеют одинаковые равновесные ситуации (либо в чистых либо в смешанных стратегиях), а их цены удовлетворяют следующему условию

$$\nu_C = \lambda \nu_A + \mu.$$

Пример 8. Элементы матриц

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix} \text{ и } C = \begin{pmatrix} 2 & 5 & 8 \\ 11 & 8 & -1 \end{pmatrix}$$

связаны равенством

$$c_{ik} = 3a_{ik} + 5, \quad i = 1, 2; \quad k = 1, 2, 3.$$

Поэтому цена игры с матрицей С легко вычисляется

$$\nu_C = 3 \cdot \nu_A + 5 = 3 \cdot 0 + 5 = 5$$

(см. пример 7).

Основные этапы поиска решения матричной игры

1-й этап – проверка наличия (или отсутствия) равновесия в чистых стратегиях (при наличии равновесной ситуации указываются соответствующие оптимальные стратегии игроков и цена игры).

2-й этап – поиск доминирующих стратегий (в случае успеха этого поиска – отбрасывание доминируемых строк и столбцов в исходной матрице игры).

3-й этап – замена игры на ее смешанное расширение и отыскание оптимальных смешанных стратегий и цены игры.

Итерационный метод решения матричных игр

Опишем метод отыскания решения матричной игры — цены игры и оптимальных смешанных стратегий, в известной степени верно отражающий некоторую реальную ситуацию накоплении опыта постепенной выработки игроками хороших стратегий в результате многих повторений конфликтных ситуаций. Основная идея этого метода заключается в том, чтобы мысленно как бы смоделировать реальное практическое «обучение» игроков в ходе самой игры, когда каждый из игроков на собственном опыте прощупывает способ поведения противника и старается отвечать на него наиболее выгодным для себя образом. Иными словами, всякий раз при возобновлении тигры игрок выбирает наиболее выгодную для себя стратегию, опираясь на предыдущий выбор противника.

Проиллюстрируем этот метод на примере игры, заданной матрицей

$$\begin{pmatrix} 2 & 0 & 3 \\ 1 & 3 & -3 \end{pmatrix}$$

(здесь maxmin = 0, minmax = 2 → седловой точки нет).

Опишем правила выбора ходов игроками, предположив, для определенности, что начинает игрок A:

ход игрока A — стратегия A_1 — $(2 \quad \mathbf{0} \quad 3)$;

игрок B выбирает свою стратегию так, чтобы выигрыш игрока A был минимален (отмечен полужирным шрифтом)

ход игрока B — стратегия B_2 — $\begin{pmatrix} 0 \\ \mathbf{3} \end{pmatrix}$

игрок A выбирает свою стратегию так, чтобы его выигрыш при стратегии B_2 игрока B был максимален (отмечен полужирным шрифтом):

ход игрока A — стратегия A_2 — $(1 \quad 3 \text{ - } 3)$;

игрок B выбирает свою стратегию так, чтобы «накопленный» выигрыш игрока A при стратегиях A_1 и A_2,

$$(2 \quad 0 \quad 3) + (1 \quad 3 \quad -3) = (3 \quad 3 \quad \mathbf{0}),$$

был минимален:

ход игрока B — стратегия B_3 — $\begin{pmatrix} 3 \\ -3 \end{pmatrix}$;

игрок A выбирает свою стратегию так, чтобы его «накопленный» выигрыш при стратегиях B_2 и B_1 игрока B,

$$\begin{pmatrix} 0 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ -3 \end{pmatrix} = \begin{pmatrix} \mathbf{3} \\ 0 \end{pmatrix}$$

был максимален:

ход игрока A — стратегия A_1 — $(2 \quad 0 \quad 3)$;

игрок B выбирает свою стратегию так, чтобы «накопленный» выигрыш игрока A при стратегиях A_1, A_2 и A_1,

$$(3 \quad 3 \quad 0) + (2 \quad 0 \quad 3) = (5 \quad \mathbf{3} \quad 3),$$

был минимален:

ход игрока B — стратегия B_2 — $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$

и т.д.

Разобьем последовательные ходы игроков A и B на пары

(*ход игрока A, ход игрока B*)

и запишем результаты в таблице

n	i	B_1	B_2	B_3	$v_*(n)$	k	A_1	A_2	$v^*(n)$	$v(n)$
1	1	2	**0**	3	0,00	2	0	**3**	3,00	1,50
2	2	3	3	**0**	0,00	3	**3**	0	1,50	0,75
3	1	5	**3**	3	1,00	2	**3**	3	1,00	1,00
4	1	7	**3**	6	0,75	2	3	**6**	1,50	1,12
5	2	8	6	**3**	0,60	3	**6**	3	1,20	0,90
6	1	10	**6**	6	1,00	2	**6**	6	1,00	1,00
7	1	12	**6**	9	0,86	2	6	**9**	1,44	1,15
8	2	13	9	**6**	0,75	3	**9**	6	1,13	0,93
9	1	15	**9**	9	1,00	2	**9**	9	1,00	1,00
10	1	17	**9**	12	0,90	2	9	**12**	1,20	1,05
11	2	18	12	**9**	0,82	3	**12**	9	1,09	0,96
12	1	20	**12**	12	1,00	2	**12**	12	1,00	1,00
…………………………………………………………………………………………..										

требующей некоторых пояснений.

Описание таблицы

1-й столбец	номер n шага (пары последовательных ходов игроков A и B)	
2-й столбец	номер i стратегии, выбранной игроком A	
3-й столбец	«накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_1 игрока B	Минимальный из этих выигрышей выделяется полужирным шрифтом
4-й столбец	«накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_2 игрока B	
5-й столбец	«накопленный» суммарный выигрыш игрока A за первые n шагов при стратегии B_3 игрока B	
6-й столбец	минимальный средний выигрыш игрока А, равный, минимальному накопленному им выигрышу за первые n шагов, деленному на число этих шагов	
7-й столбец	номер k стратегии, выбранной игроком B	
8-й столбец	«накопленный» суммарный выигрыш игрока А за первые n шагов при стратегии A_1	Максимальный из этих выигрышей выделяется полужирным шрифтом
9-й столбец	«накопленный» суммарный выигрыш игрока А за первые n шагов при стратегии A_2	
10-й столбец	максимальный средний выигрыш игрока A, равный максимальному накопленному им выигрышу за первые n шагов, деленному на число этих шагов	
11-й столбец	среднее арифметическое минимального среднего выигрыша и максимального среднего выигрыша игрока A	

Решение игры определяется приближенно по окончании любого из шагов.

Например, за приближенную цену игры можно взять среднее арифметическое $v(n)$, полученное на n-м шаге. Смешанные стратегии противников определяются частотами появления чистых стратегий.

После 9-го шага имеем

$$v(9) = 1{,}00.$$

При этом игрок A 6 раз использовал стратегию A_1 и 3 раза стратегию A_2. В свое очередь игрок B 6 раз применял стратегию B_2, 3 раза стратегию B_3, а стратегией B_1 не пользовался вообще. Отсюда получаем, что

$$P_9 = \left\{\frac{6}{9},\ \frac{3}{9}\right\} \approx \{0{,}67;\ 0{,}33\},\quad Q_9 = \left\{0,\ \frac{6}{9},\ \frac{3}{9}\right\} \approx \{0;\ 0{,}67;\ 0{,}33\}$$

Соответственно, после 10-го шага получаем —

$$\nu(10) = 1{,}05,\quad P_{10} = \left\{\frac{7}{10},\ \frac{3}{10}\right\} = \{0{,}7;\ 0{,}3\},\quad Q_{10} = \left\{0,\ \frac{7}{10},\ \frac{3}{10}\right\} = \{0;\ 0{,}7;\ 0{,}3\}.$$

Данная игра легко решается графически. Вот точный ответ:

$$\nu = 1,\quad P = \left\{\frac{2}{3},\ \frac{1}{3}\right\},\quad Q = \left\{0,\ \frac{2}{3},\ \frac{1}{3}\right\}.$$

Сравнивая результаты, полученные на 9-м, 10-м, а также 11-м и 12-м шагах:

$$\begin{array}{ll} \nu(11) = 0{,}96 & \nu(11) = 1{,}00 \\ P_{11} = \left\{\frac{7}{11}, \frac{4}{11}\right\} \approx \{0{,}64; 0{,}36\} & P_{12} = \left\{\frac{8}{12}, \frac{4}{12}\right\} \approx \{0{,}67; 0{,}33\} \\ Q_{11} = \left\{0, \frac{7}{11}, \frac{4}{11}\right\} \approx \{0; 0{,}64; 0{,}36\} & Q_{12} = \left\{0, \frac{8}{12}, \frac{4}{12}\right\} \approx \{0; 0{,}67; 0{,}33\} \end{array}$$

замечаем, что по мере увеличения числа шагов значения все меньше отличаются от точных.

Сделаем несколько замечаний.

Замечание 1. При увеличении числа шагов все три величины $\nu_*(n)$, $\nu^*(n)$ и $\nu(n)$ будут приближаться к цене игры ν, но среднее арифметическое $\nu(n)$ будет приближаться к ν сравнительно быстрее.

Замечание 2. Хотя сходимость итераций весьма медленна, тем не менее, даже такой небольшой расчет всегда дает возможность находить ориентировочное значение цены игры и доли чистых стратегий.

Замечание 3. Сравнительно медленная скорость сходимости можно объяснить целым рядом причин. Укажем одну из них, психологически наиболее интересную. Если, к примеру, игрок A уже получил оптимальную смешанную стратегию, то он не склонен останавливаться на ней. Отнюдь нет — он продолжит попытки выиграть у противника B побольше, особенно если последний еще не достиг оптимальной смешанной стратегии. Тем самым, игрок A может невольно ухудшить свое положение.

Замечание 4. Отметим два основных преимущества описанного метода:
1) итерационный метод прост и одновременно универсален (при его помощи можно легко найти приближенное решение любой матричной игры),
2) объем и сложность вычислений сравнительно слабо растут по мере увеличения числа стратегий игроков (размеров матрицы игры).

Сведение матричной игры к задаче линейного программирования

Рассмотрим m x n игру с платежной матрицей

$$A = (a_{ik}).$$

Без ограничения общности будем считать, что все элементы матрицы A положительны (этого всегда можно добиться, пользуясь аффинным правилом, преобразующим заданную матрицу игры, но не изменяющим оптимальных смешанных стратегий игроков). Тем самым, искомая цена игры ν — положительное число.

Интересы игрока А

Из теоремы о свойствах оптимальных смешанных стратегий игроков вытекает, что при любой чистой стратегии B_k игрока B, $k = 1, 2, \ldots, n$, оптимальная смешанная стратегия P

= $\{p_1, p_2, \dots, p_m\}$ игрока A обеспечивает его средний выигрыш, не меньший v. Иными словами, выполняются соотношения

$$\sum_{i=1}^{m} a_{ik} p_i \ge v,\ k = 1, 2, \dots, n, \qquad \sum_{i=1}^{m} p_i = 1,\ p_i \ge 0,\ \ i = 1, 2, \dots, m,$$

которые с учетом обозначений

$$x_i = \frac{p_i}{v},\ \ i = 1, 2, \dots, m,$$

можно записать так

$$\sum_{i=1}^{m} a_{ik} x_i \ge v,\ k = 1, 2, \dots, n, \qquad \sum_{i=1}^{m} x_i = \frac{1}{v},\ x_i \ge 0,\ \ i = 1, 2, \dots, m.$$

Поскольку игрок A стремится сделать гарантированный выигрыш максимально возможным, то задача отыскания решения матричной игры сводится к следующей задаче: найти неотрицательные величины $x_1, x_2, \dots, x_m$, удовлетворяющие неравенствам

$$\sum_{i=1}^{m} a_{ik} x_i \ge 1,\ k = 1, 2, \dots, n,$$

и такие, что их сумма минимальна

$$\sum_{i=1}^{m} x_i \to \min.$$

Интересы игрока В

Аналогичным образом заключаем, что оптимальная смешанная стратегия $Q = \{q_1, q_2, \dots, q_n\}$ игрока B при любой чистой Стратегии A_1 игрока A, $i = 1, 2, \dots, m$, обеспечивает его средний проигрыш, не больший v. Иными словами, выполняются соотношения

$$\sum_{k=1}^{n} a_{ik} q_k \le v,\ i = 1, 2, \dots, m, \qquad \sum_{k=1}^{n} q_k = 1,\ q_k \ge 0,\ \ k = 1, 2, \dots, n.$$

которые с учетом обозначений

$$y_k = \frac{q_k}{v},\ \ k = 1, 2, \dots, n,$$

можно записать так

$$\sum_{k=1}^{n} a_{ik} y_k \le 1,\ i = 1, 2, \dots, m, \qquad \sum_{k=1}^{n} y_k = \frac{1}{v},\ y_k \ge 0,\ \ k = 1, 2, \dots, n.$$

Поскольку игрок B стремится сделать свой гарантированный проигрыш минимально возможным, то задача отыскания решения матричной игры сводится к следующей задаче:

найти неотрицательные величины $y_1, y_2, \dots, y_n$, удовлетворяющие неравенствам

$$\sum_{k=1}^{n} a_{ik} y_k \le 1,\ i = 1, 2, \dots, m,$$

и такие, что их сумма максимальна

$$\sum_{k=1}^{n} y_k \to \max$$

Тем самым, мы получаем следующий важный результат.

Теорема 3. Решение матричной игры с положительной платежной матрицей (a_{ik}) равносильно решению двойственных задач линейного программирования

(А) $$\sum_{i=1}^{m} x_i \to \min, \qquad \sum_{i=1}^{m} a_{ik} x_i \ge 1,\ k = 1, 2, \dots, n,\ x_i \ge 0,\ \ i = 1, 2, \dots, m,$$

(В) $$\sum_{k=1}^{n} y_k \to \max, \qquad \sum_{k=1}^{n} a_{ik} y_k \le 1,\ i = 1, 2, \dots, m,\ y_k \ge 0,\ \ k = 1, 2, \dots, n.$$

При этом цена игры

$$\nu = \frac{1}{\Theta},$$

где Θ — величина, обратная общему значению оптимальных сумм,

$$\Theta = \sum_{i=1}^{m} x_i^0 = \sum_{k=1}^{n} y_k^0$$

а оптимальные значения p_i^0 и q_k^0 связаны с оптимальными x_i^0 и y_k^0 посредством равенств

$$p_i^0 = \frac{x_i^0}{\Theta}, \quad i = 1, 2, \dots, m, \qquad q_k^0 = \frac{y_k^0}{\Theta}, \quad k = 1, 2, \dots, n.$$

Алгоритм решения матричной игры

1-й шаг. Ко всем элементам исходной матрицы игры прибавляется одно и то же положительное число γ так, чтобы все элементы новой матрицы были строго положительны.

2-й шаг. Решаются двойственные задачи линейного программирования (*A*) и (*B*) (например, симплекс-методом, или как-нибудь иначе). Находятся наборы x_i^0, y_k^0 и число Θ.

3-й шаг. Строятся оптимальные смешанные стратегии игроков *A* и *B* соответственно

$$p_i^0 = \frac{x_i^0}{\Theta}, \qquad q_k^0 = \frac{y_k^0}{\Theta}.$$

4-й шаг. Вычисляется цена игры

$$\nu_A = \frac{1}{\Theta} - \gamma$$

Пример 9. Рассмотрим 2 x 2 игру с матрицей

$$\begin{pmatrix} 5 & 1 \\ 3 & 4 \end{pmatrix}$$

Соответствующие ей задачи линейного программирования имеют вид

$$\begin{array}{ll} x_1 + x_2 \to \min & y_1 + y_2 \to \max \\ 5x_1 + 3x_2 \ge 1, \quad x_1 \ge 0 & 5y_1 + y_2 \le 1, \quad y_1 \ge 0 \\ x_1 + 4x_2 \ge 1, \quad x_2 \ge 0 & 3y_1 + 4y_2 \le 1, \quad y_2 \ge 0 \end{array}$$

Решение.

1-й шаг. Все элементы платежной матрицы положительны.

2-й шаг. Строим решения обеих задач линейного программирования, пользуясь графическим методом. В результате получаем, что

$$x_1^0 = \frac{1}{17}, \quad x_2^0 = \frac{4}{17}, \quad y_1^0 = \frac{3}{17}, \quad y_2^0 = \frac{2}{17}, \quad \Theta = \frac{5}{17}$$

3-й шаг.

$$p_1^0 = \frac{1}{5}, \quad p_2^0 = \frac{4}{5}, \quad q_1^0 = \frac{3}{5}, \quad q_2^0 = \frac{2}{5}.$$

4-й шаг.

$$\nu_A = \frac{17}{5} = 3\frac{2}{5}$$

4. Примеры задач, сводимых к матричным играм

В чистом виде антагонистические конфликты встречаются редко (разве только в боевых действиях и в спортивных состязаниях). Однако довольно часто конфликты» в которых интересы сторон противоположны, при допущении, что множество способов действия сторон конечно, можно моделировать матричными играми.

Рассмотрим несколько конкретных ситуаций.

Пример 10. «Планирование поема». Сельскохозяйственное предприятие имеет возможность выращивать две культуры — A_1 и A_2. Необходимо определить, как сеять эти культуры, если при прочих равных, условиях их урожаи зависят от погоды, а план посева должен обеспечить наибольший доход (прибыль от реализации выращенной культуры определяется полученным объемом). В зоне рискованного земледелия (а таковой является большая часть России) планирование посева должно осуществляться с учетом наименее благоприятного состояния погоды.

Таким образом, одной из сторон выступает сельскохозяйственное предприятие, заинтересованное в том, чтобы получить наибольший доход (игрок A), а другой стороной — природа, способная навредить сельскохозяйственному предприятию в максимальной степени (от нее зависят погодные условия) и преследующая тем самым прямо противоположные цели (игрок B).

Принятие природы за противника равносильно планированию посева с учетом наиболее неблагоприятных условий; если же погодные условия окажутся благоприятными, то выбранный план даст возможность увеличить доход.

Налицо антагонистический конфликт, в котором у игрока A две стратегии — A_1 и A_2, а у игрока B три — B_1 (засушливое лето), B_2 (нормальное лето) и B_3 (дождливое лето).

В качестве выигрыша игрока A возьмем прибыль от реализации и будем считать, что расчеты прибыли сельскохозяйственного предприятия (в млрд руб.) в зависимости от состояний погоды сведены в следующую матрицу

$$\begin{pmatrix} 8 & 5 & 3 \\ 2 & 3 & 6 \end{pmatrix}.$$

Нетрудно заметить, что седловой точки у этой матрицы нет. Поэтому оптимальная стратегия игрока A будет смешанной. Применяя графический метод, получаем

$$P = \left\{\frac{3}{5}, \frac{2}{5}\right\}, \quad v = 4\frac{1}{5}.$$

Замечание. Здесь мы столкнулись со сравнительно редкой ситуацией, когда оптимальная смешанная стратегия одного из игроков допускает так называемую «физическую» реализацию. Полученное решение сельскохозяйственное предприятие может использовать так:

на $\frac{3}{5}$ всех площадей выращивать культуру A_1,

на $\frac{2}{5}$ всех площадей выращивать культуру A_2

и получать прибыль в размере, не меньшем $4\frac{1}{5}$ млрд руб.

Пример 11. «Переговоры о заключении контракта между профсоюзом и администрацией». Рассмотрим фирму, администрация которой ведет переговоры с профсоюзом рабочих и служащих о заключении контракта.

Предположим, что платежная матрица, отражающая интересы договаривающихся сторон, имеет следующий вид

$$\begin{pmatrix} 75 & 105 & 65 & 45 \\ 70 & 60 & 55 & 40 \\ 80 & 90 & 35 & 50 \\ 95 & 100 & 50 & 55 \end{pmatrix}$$

Выплаты указаны в центах в час и представляют собой среднюю зарплату служащего фирмы вместе со всеми добавками. Тем самым, заданная матрица описывает прибыль профсоюза (игрок A) и затраты администрации фирмы (игрок B).

Ясно, что профсоюз стремится максимизировать доходы рабочих и служащих, в то время как администрации хотелось бы минимизировать собственные потери.

Нетрудно заметить, что седловой точки у платежной матрицы нет. Кроме того, для дальнейшего анализа существенными являются лишь стратегии A_1 и A_4 игрока A и стратегии B_3 и B_4 игрока B (в этом нетрудно убедиться, воспользовавшись правилом доминирования стратегий). В результате соответствующего усечения получим матрицу

$$\begin{pmatrix} 65 & 45 \\ 50 & 55 \end{pmatrix}.$$

Элементы матрицы

$$\begin{pmatrix} 4 & 0 \\ 1 & 2 \end{pmatrix}$$

связаны с элементами предыдущей матрицы соотношениями

65 + 5 x 4 + 45,   45 = 5 x 4 + 45,   50 = 5 x 1 + 45,   55 = 5 x 2 +45.

Воспользовавшись графическим методом, в итоге получим

$$P = \left\{\frac{1}{5}, 0, 0, \frac{4}{5}\right\}, \quad P = \left\{0, 0, \frac{2}{5}, \frac{3}{5}\right\}, \quad v = 53.$$

Тем самым, профсоюзу следует выбирать стратегию A_1 в 20 % случаев и стратегию A_4 в 80 %. Что касается администрации, то ей следует выбирать стратегию B_3 с вероятностью 0,4 и стратегию B_4 с вероятностью 0,6. При этом ожидаемая цена игры равна 53.

Замечание. Следует отметить, что если процесс переговоров будет повторяться много раз, то среднему должно сходиться к ожидаемому значению 53. Если же переговоры пройдут лишь единожды, то реальный результат подучится при выборе каждым игроком некоторой своей чистой стратегии. Поэтому один из игроков, профсоюз или администрация, будет неудовлетворен.

Пример 12. «Локальный конфликт». Рассмотрим войну между двумя небольшими государствами A и B, которая ведется в течение 30 дней.

Для бомбардировки небольшого моста — важного военного объекта страны B — страна A использует оба имеющихся у нее самолета. Разрушенный мост восстанавливается в течение суток, а каждый самолет совершает один полет в день по одному из двух воздушных маршрутов, соединяющих эти страны. У страны B есть два зенитных орудия, при помощи которых можно сбивать самолеты страны A. Если самолет собьют, то некая третья страна в течение суток поставит стране A новый самолет.

Страна A может послать самолеты либо по одному маршруту, либо по разным. Страна B может поместить либо обе зенитки на одном маршруте, либо по одной зенитке на каждый маршрут.

Если один самолет летит по маршруту, на котором расположена одна зенитка, то этот самолет будет сбит. Если два самолета летят по маршруту, на котором расположены две зенитки, то оба самолета будут сбиты. Если два самолета летят по маршруту, на котором расположена одна зенитка, то сбит будет только один самолет. Если самолет доберется до цели, то мост будет уничтожен.

У страны A есть две стратегии:
послать самолеты по разным маршрутам — A_1,
послать самолеты по одному маршруту — A_2.
У страны B — также две стратегии:
поместить зенитки на разных маршрутах — B_1,

поместить зенитки на одном маршруте — B_2.

Если страна A выберет стратегию A_1, а страна В — стратегию B_1, то страна A получит нулевой выигрыш, так как ни один из самолетов не достигнет цели.

Если страна A выберет стратегию A_2, а страна B — стратегию B_1, то хота бы один самолет достигнет цели и вероятность разрушения моста будет равна 1.

Если страна A выберет стратегию A_1, а страна B — стратегию B_2, то вновь хотя бы один самолет достигнет цели и вероятность разрушения моста будет равна 1.

Если страна A выберет стратегию A_2, а страна B — стратегию B_2, то страна A с вероятностью 1/2 выберет маршрут, на котором установлены зенитки, и, следовательно, цель будет уничтожена с вероятностью 1/2.

Оформим результаты проведенного анализа в стандартной игровой форме:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1/2 \end{pmatrix}$$

При помощи графического метода получаем оптимальные смешанные стратегии игроков и цену игры

$$P = \left\{\frac{1}{3},\ \frac{2}{3}\right\},\quad Q = \left\{\frac{1}{3},\ \frac{2}{3}\right\},\quad v = \frac{2}{3}$$

Это означает, что если страна A будет посылать самолеты по разным маршрутам в течение десяти дней из тридцати, отпущенных на войну (и, значит, по одному маршруту в течение двадцати дней), то в среднем страна A будет иметь 66,7 % удачных случаев (мост будем находиться в нерабочем состоянии). Воспользовавшись для своих зениток предложенным выбором, страна B не позволит бомбить мост чаще, чем в 66,7 % случаев.

Несколько слов в заключение

Матричные игры моделируют конфликтные ситуации, в которых каждая из сторон-участниц делает свой ход одновременно со второй стороной. При этом наибольший интерес представляет случай, когда игра не заканчивается сразу же после совершения игроками одной такой пары одновременных ходов, а повторяется многократно. Причем считается, что перед каждым возобновлением игры игроки не получают никаких новых сведений ни о конфликте, ни о возможных действиях противной стороны. Иными словами, при многократном повторении матричной игры каждая из сторон всякий раз оказывается перед выбором некоторой стратегии из одного и того же множества стратегий, неизменного у каждого из игроков.

Тем не менее, в таких многократно повторяющихся обстоятельствах большую роль играет анализ игры, как предварительный, так и промежуточный.

В результате разумно проведенного предварительного анализа матричной игры заинтересованная в анализе сторона может определить свою линию поведения (правило выбора стратегий) на всю серию игр. Разумеется, описанный нами выше максиминный подход является далеко не единственным средством. Однако не следует забывать, что принципиальной особенностью этого подхода является то обстоятельство, что игрок, придерживающийся выводимого на его основе правила выбора стратегий, заранее может довольно точно оценить нетривиальные размеры своего гарантированного выигрыша. Кроме того, максиминный подход позволяет сводить задачу поиска решения игры к рассмотрению сравнительно несложных задач линейного программирования и, тем самым, получать эффективные рекомендации по тому, как лучше выбирать стратегии в конкретной игре при многократном ее повторении.

Если игра повторяется много раз, то некоторые дополнительные сведения — какие именно стратегии выбирает противная сторона и какими правилами выбора стратегий она руководствуется — игрок все же получает. На основании этих сведений и результатов предварительного анализа игры он может довольно точно оценить противника и, если тот не придерживается компромиссного максиминного подхода, внести соответствующие изменения в собственную линию поведения и увеличить выигрыш.

6. О классификации игр

Реальные конфликтные ситуации приводят к различным видам игр. К настоящему времени общепризнанной классификации игр пока не сложилось. Тем не менее, легко заметить, что игры различаются по целому ряду признаков: по количеству участвующих в них игроков, по количеству возможных стратегий, по характеру взаимоотношений между игроками, по характеру выигрышей, по виду функций выигрышей, по количеству ходов, по характеру информационной обеспеченности игроков и т. д.

Остановимся на этих различиях чуть подробнее.

В зависимости от количества игроков определяют игры трех типов: игры *одного игрока* (в теории игр, как правило, не рассматриваются), игры *двух игроков* (наиболее изученный класс игр) и игры *n игроков* (успехов в изучении которых сравнительно немного вследствие возникающих принципиальных трудностей).

По количеству стратегий игроков игры делятся на конечные (каждый из игроков имеет конечное число возможных стратегий) и бесконечные (где хотя бы один из игроков имеет бесконечное количество возможных стратегий).

По характеру взаимоотношений между игроками игры делятся на *бескоалиционные* (в которых игроки не имеют права вступать в соглашения и образовывать коалиции), *коалиционные* (в которых игроки могут вступать в соглашения и образовывать коалиции) и *кооперативные* (в которых соглашения, связывающие игроков, определены наперед и обязательны):

По характеру выигрышей различают игры с *нулевой суммой* (общий капитал игроков не изменяется, а просто перераспределяется между игроками в зависимости от получающихся исходов) и игры с *ненулевой суммой*.

Игру двух игроков с нулевой суммой часто называют *антагонистической*, вследствие того, что цели игроков в них прямо противоположны: выигрыш одного из игроков происходит только за счет проигрыша другого.

По виду функций выигрышей игры делятся на: *матричные* игры (рассмотренные выше), *биматричные* игры, игры *типа дуэлей*, *непрерывные* игры, *выпуклые* игры и др.

По количеству ходов игры делятся на *одношаговые* (завершающиеся после одного хода каждого из игроков) и *многошаговые*, которые, в свою очередь, делятся на позиционные игры (каждый из игроков может последовательно во времени делать несколько ходов), *стохастические* игры (где при выборе новых позиций имеется определенная вероятность возврата на предшествующую позицию), *дифференциальные* игры (в которых допускается делать ходы непрерывно и подчинять поведение игроков условиями, описываемыми дифференциальными уравнениями), игры *типа дуэлей* (характеризующиеся моментом выбора хода и вероятностями получения выигрышей в зависимости от времени, прошедшего от начала игры до момента выбора).

По характеру информационной обеспеченности игроков различают игры с *полной информацией* (на каждом ходе игры каждому игроку известно, какие выборы были сделаны ранее всеми игроками) и игры с *неполной информацией* (если в игре не все известно о предыдущих выборах).

Существуют, разумеется, и другие виды игр.

В зависимости от вида игры разрабатывается и метод ее решения. Однако читатель, видимо, уже заметил, что структура предложенного выше описания основных направлений различения игр не является древовидной: одни и те же виды игр оказываются в разных классах.

Впрочем, возможны и иные подходы к разбиению игр.

Далее мы рассмотрим некоторые из указанных видов игр, а именно позиционные и биматричные игры.

Часть II ПОЗИЦИОННЫЕ ИГРЫ

Во многих практически важных конфликтных ситуациях, располагая той или иной информацией об их прошлом развитии, стороны-участницы совершают свой выбор не раз и навсегда, а последовательно во времени, шаг за шагом. Тем самым, они используют стратегии, отражающие как динамику конфликта, так и степень собственной информированности о фактически складывающейся обстановке в развитии этого конфликта.

1. Структура позиционной игры

Одним из классов игр, описывающих конфликты, динамика которых оказывает влияние на поведение участников, являются так называемые позиционные игры.

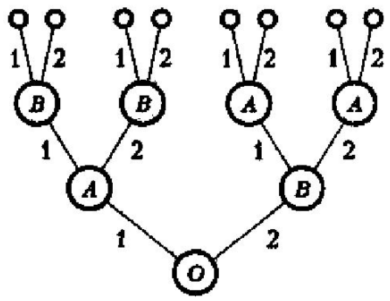
Позиционная игра — это бескоалиционная игра, моделирующая процессы последовательного принятия решений игроками в условиях меняющейся во времени и, вообще говоря, неполной информации.

Процесс самой игры состоит в последовательном переходе от одного состояния игры к другому состоянию, который осуществляется либо путем выбора игроками одного из возможных действий в соответствии с правилами игры, либо случайным образом (*случайный ход*).

В качестве примеров позиционных игр можно привести крестики-нолики, шашки, шахматы, карточные игры, домино и др. Интересно, что право выбора первого хода в этих играх часто определяется случайным образом.

Состояния игры принято называть позициями (отсюда и название — позиционные игры), а возможные выборы в каждой позиции — альтернативами.

Характерной особенностью позиционной игры является возможность представления множества позиций в виде древовидного упорядоченного множества, которое называется *деревом игры* (рис. 1).



Для определенности мы будем рассматривать позиционные игры, в каждой позиции которых, кроме окончательных, ровно две альтернативы — первая и вторая.

Замечание. Символ *O, A* или *B* в кружке указывает, кто из игроков (*O, A* или *B*) делает очередной ход в заданной позиции. При этом символом *O* обычно обозначается ход в игре, осуществляемый не игроком, а каким-нибудь случайным механизмом (иногда его называют *природой*). Например, в позиционной игре, представленной на рис. 1 своим деревом, первый ход производится случайно.

Пользуясь графическим описанием игры, можно сказать, что процесс игры состоит в переходе от начальной позиции к окончательной через непосредственно следующие одна за другой промежуточные позиции.

Каждая окончательная вершина определяет единственную цепь (последовательность идущих друг за другом звеньев), связывающую начальную вершину с данной (рис. 2). Такая цепь называется *партией*. На рис. 2 одна из партий выделена жирными линиями. Число различных партий равно числу окончательных вершин (позиций).

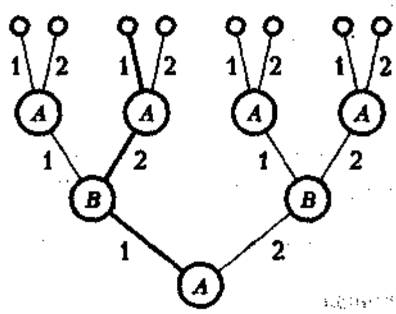


Рис. 2

В каждой окончательной позиции задан числовой выигрыш игрока *A*.

Замечание. Мы будем рассматривать здесь только антагонистические позиционные игры.

В шахматах функция выигрышей игрока *A* (белых) определяется так:

+1 на выигрываемых партиях,

0 на ничейных партиях,

-1 на проигрываемых партиях.

Функция выигрышей игрока *B* (черных) отличается от функции выигрышей белых только знаком.

Различают позиционные игры с *полной информацией* и позиционные игры с *неполной информацией*.

В позиционных играх с полной информацией (пример — шашки, шахматы) каждый игрок при своем ходе знает ту позицию дерева игры, в которой он находится.

В позиционных играх с неполной информацией (пример — домино) игрок, де лающий ход, не знает точно, в какой именно позиции дерева игры он фактически находится. Этому игроку известно лишь некоторое множество позиций, включающее в себя его фактическую позицию. Такое множество позиций называется информационным множеством.

Таким образом, в игре с неполной информацией игрок при своем ходе знает, в каком информационном множестве он находится, но ему неизвестно, в какой именно позиции этого множества.

Позиции, принадлежащие одному и тому же информационному множеству, объединяются пунктирными линиями.

Рассмотрим примеры двух игр, состоящих из двух ходов, которые последовательно делают участвующие в ней игроки *A* и *B*. Начинает игрок *A*: он выбирает одну из двух возможных альтернатив — число x, равное либо 1 (первая альтернатива), либо 2 (вторая альтернатива). На ход игрока *A* игрок *B* отвечает своим ходом, выбирая одну из двух возможных альтернатив — число y, равное либо 1 (первая альтернатива), либо 2 (вторая альтернатива).

И в результате игрок *A* получает вознаграждение или вынужден платить штраф.

Пример 13.

1-й ход. Игрок *A* выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход. Игрок *B* выбирает число y из множества двух чисел $\{1,2\}$, зная выбор числа x игроком *A*.

Функция $W(x, y)$ выплат игроку *A* за счет игрока *B* задается так

$$W(1, 1) = 1, \quad W(2, 1) = -2,$$
$$W(1, 2) = -1, \quad W(2, 2) = 2.$$

На рис. 3 показаны дерево игры и информационные множества.

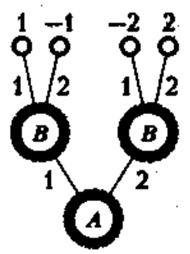


Рис. 3

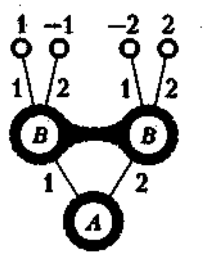


Рис.4

Пример 14. В случав, вели выполнены все условия предыдущего примера, кроме одного — хода игрока B.

2-й ход — игрок B выбирает число у из множества двух чисел $\{1, 2\}$, не зная выбора числа x игроком A, информационные множества выглядят так, как показано на рис. 4.

2. Нормализация позиционной игры

Заранее определенную последовательность ходов игрока, выбранную им в зависимости от информации о ходах другого игрока и ходах игрока O (природы), будем называть *чистой стратегией* этого игрока.

В том случае, если в игре нет случайных ходов (игрок O в игре не участвует), выбор игроком A и игроком B чистых стратегий однозначно определяет исход игры — приводит к окончательной позиции, где игрок A. и получает свой выигрыш. Это обстоятельство позволяет сводить позиционную игру к матричной игре.

Процесс сведения позиционной игры к матричной называется *нормализацией позиционной* игры.

Покажем на нескольких примерах, как это делается.

Пример 13 (продолжение). Опишем стратегии игроков.

Стратегию игрока A можно задать одним числом x, показывающим, какую альтернативу, первую или вторую, выбрал игрок.

Тем самым, у игрока A две чистых стратегии:

$$A_1 \text{ —выбрать } x = 1, \quad A_2 \text{ — выбрать } x = 2.$$

Стратегию игрока B, принимая во внимание, что выбор игрока A на 1-м ходе ему известен, удобно описывать упорядоченной парой

$$[y_1, y_2].$$

Здесь y_1 (y_1 = 1, 2) — альтернатива, выбираемая игроком B при условии, что игрок A выбрал первую альтернативу, x = 1, а y_2 (y_2 = 1, 2) — альтернатива, выбираемая игроком B при условии, что игрок A выбрал вторую альтернативу, x = 2.

Например, выбор игроком B стратегии [2, 1] означает, что если на 1-м ходе игрок A выбрал $x = 1$, то игрок B на своем ходе должен выбрать $y = 2$. Если же на 1-м ходе игрок A выбрал $x = 2$, то согласно этой стратегии игрок B на своем ходе должен выбрать $y = 1$.

Таким образом, у игрока B четыре чистых стратегии:

$$B_1 \text{ — } [1,1], \quad y = 1 \quad \text{при любом выборе } x;$$
$$B_2 \text{ — } [1, 2], \quad y = x \quad \text{при любом выборе } x;$$
$$B_3 \text{ — } [2,1], \quad y \neq x \quad \text{при любом выборе } x;$$
$$B_4 \text{ — } [2, 2], \quad y = 2 \quad \text{при любом выборе } x.$$

Покажем теперь, как рассчитываются выигрыши игрока A в зависимости от примененных стратегий.

Пусть, например, игрок A выбрал стратегию A_1 — (1), а игрок B — стратегию B_2 — [1, 2]. Тогда $x = 1$, а из стратегии [1, 2] вытекает, что $y = 1$. Отсюда

$$W(x, y) = W(1, 1) = 1.$$

Остальные выигрыши рассчитываются совершенно аналогично.

Результаты расчетов записывается обычно или в виде таблицы выигрышей игрока A

B_1	B_2	B_3	B_4

		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	$x = 1$	$W(1, 1)$	$W(1, 1)$	$W(1, 2)$	$W(1, 2)$
A_2	$x = 2$	$W(2, 1)$	$W(2, 2)$	$W(2, 1)$	$W(2, 2)$

или в вида матрицы игры

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ -2 & 2 & -2 & 2 \end{pmatrix},$$

где, как обычно, строки соответствуют стратегиям игрока A, а столбцы — стратегиям игрока B,

Полученная матрица имеет седловую точку. Оптимальные стратегии игроков: A_1 — (1) и B_3 — [2, 1]. Тем самым, игрок A на 1-м ходе выбирает $x = 1$, а игрок B на 2-м ходе выбирает $y = 2$. Цена игры $v =$ -1.

Пример 14 (продолжение). Опишем стратегии игроков.

У игрока A они те же, что и в предыдущем примере

A_1 – выбрать $x = 1$, A_2 — выбрать $x = 2$.

Так как игроку B выбор игрока A неизвестен, то есть игрок B не знает, в какой именно из двух позиций он находится (см. рис. 4), то у него те же две стратегии:

B_1 — выбрать $y = 1$, B_2 — выбрать $y = 2$.

Соответствующие таблица выигрышей игрока A и матрица игры имеют следующий вид

		B_1	B_2
		$Y = 1$	$y = 2$
A_1	$x = 1$	$W(1, 1)$	$W(1, 2)$
A_2	$x = 2$	$W(2, 1)$	$W(2, 2)$

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}$$

Полученная матрица седловой точки не имеет. Оптимальные смешанные стратегии игроков: $P = \{2/3, 1/3\}$ и $Q = \{1/2, 1/2\}$. Цена игры $v = 0$.

Замечание 1. На этих двух примерах хорошо видно, что результат сведения позиционной игры к матричной напрямую зависит от степени информированности игроков. В частности, отсутствие у игрока B сведений о выборе, сделанном игроком A приводит к уменьшению количества его возможных стратегий. Сравнивая, ответы, полученные в примерах 13 и 14, замечаем, что, снижение уровня информированности игрока (в данном случае — игрока B) делает для него исход игры менее, благоприятными.

Замечание 2. Приведенные выше примеры не исчерпывают всех возможных вариантов даже в этом, самом простом, случае двухходовых позиционных игр.

Рассмотрим теперь несколько примеров сведения к матричным играм позиционных игр, состоящих из трех ходов, сосредоточив при этом основное внимание на одном из наиболее ответственных шагов нормализации — описании стратегий игроков.

Пример 15.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: зная выбранное игроком A число x, он выбирает число у из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: не зная о выбранном игроке B числе у на 2-м ходе и забыв выбранное им самим на 1-м ходе число x, он выбирает число z из множества двух чисел $\{1, 2\}$.

После этого игрок A получает вознаграждение $W(x, y, z)$ за счет игрока B, например, такое:

$W(1, 1, 1) = -2$ $\quad\quad$ $W(2, 1, 1) = 3$

$$W(1, 1, 2) = 4 \qquad W(2, 1, 2) = 0$$
$$W(1, 2, 1) = 1 \qquad W(2, 2, 1) = -3$$
$$W(1, 2, 2) = -4 \qquad W(2, 2, 2) = 5$$

На рис.5 показаны дерево игры и информационные множества.

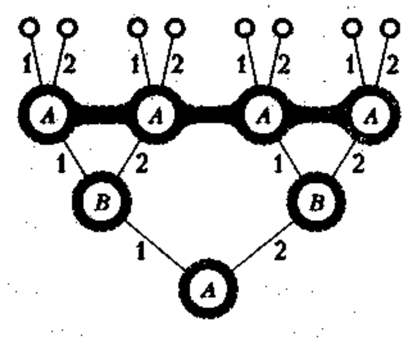


Рис. 5

Нормализуем эту игру.

Поскольку игроку B выбор игрока A на 1-м ходе известен, то у игрока B те же четыре стратегии, что и в примере 13:

$$B_1 – [1, 1], B_2 – [1, 2], B_3 – [2, 1], B_4 – [2, 2].$$

Игрок A на 3-м ходе не знает предыдущих выборов — ни значения x, ни значения y. Поэтому каждая его стратегия состоит просто из пары чисел (x, z), где x ($x = 1, 2$) — альтернатива, выбираемая игроком A на 1-м ходе, а z ($z = 1, 2$) — альтернатива, выбираемая игроком A на 3-м ходе.

Например, выбор игроком A стратегии (2, 1) означает, что на 1-м ходе он выбирает $x = 2$, а на 3-м ходе — $z = 1$.

Таким образом, у игрока А четыре стратегии:

$$A_1 – [1, 1], A_2 – [1, 2], A_3 – [2, 1], A_4 – [2, 2].$$

Покажем теперь, как рассчитываются выигрыши игрока A в зависимости от стратегий, применяемых в данной игре. Пусть, например, игрок A выбрал стратегию A_2 — (1, 2), а игрок B — стратегию B_3 — (2, 1]• Тогда $x = 1$, откуда вытекает, что $y = 2$. Значение $z = 2$ выбрано игроком A независимо от выбора игрока B. Вычисляя значение функции выигрышей для этого набора, получаем

$$W(x, y, z) = W(1, 2, 2) = -4.$$

В результате подобных рассуждений получаются и остальные пятнадцать выигрышей. Это позволяет построить таблицу выигрышей игрока A. Имеем

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1, 1)	W(1, 1, 1)	W(1, 1, 1)	W(1, 2, 1)	W(1, 2, 1)
A_2	(1, 2)	W(1, 1, 2)	W(1, 1, 2)	W(1, 2, 2)	W(1, 2, 2)
A_3	(2, 1)	W(2, 1, 1)	W(2, 2, 1)	W(2, 1, 1)	W(2, 2, 1)
A_4	(2, 2)	W(2, 1, 2)	W(2, 2, 2)	W(2, 1, 2)	W(2, 2, 2)

или

$$\begin{pmatrix} -2 & -2 & 1 & 1 \\ 4 & 4 & -4 & -4 \\ 3 & -3 & 3 & -3 \\ 0 & 5 & 0 & 5 \end{pmatrix}.$$

Пример 16.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: не зная о выборе игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, не зная ни значения x, ни значения y.

После этого игроки расплачиваются по правилу, указанному в примере 15.

Графическое представление этой игры показано на рис. 6.

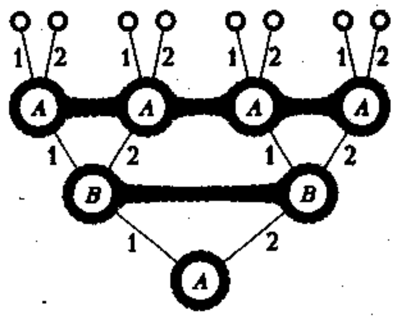


Рис. 6

Ясно, что у игрока A те же четыре стратегии, что и в примере 15:

$$A_1 - [1, 1], A_2 - [1, 2], A_3 - [2, 1], A_4 - [2, 2].$$

У игрока B всего две стратегии:

$$B_1 - \text{выбрать } y = 1,\ B_2 - \text{выбрать } y = 2$$

В этом случае (весьма слабой информированности игроков) таблица выигрышей игрока A и соответствующая матрица строятся совсем просто. Имеем

		B_1	B_2
		$y = 1$	$y = 2$
A_1	(1, 1)	$W(1, 1, 1)$	$W(1, 1, 1)$
A_2	(1, 2)	$W(1, 1, 2)$	$W(1, 2, 2)$
A_3	(2, 1)	$W(2, 1, 1)$	$W(2, 2, 1)$
A_4	(2, 2)	$W(2, 1, 2)$	$W(2, 2, 2)$

$$\begin{pmatrix} -2 & 1 \\ 4 & -4 \\ 3 & -3 \\ 0 & 5 \end{pmatrix}$$

Оптимальные смешанные стратегии игроков и цена игры соответственно равны:

$$P = \left\{0, \frac{9}{13}, 0, \frac{4}{13}\right\}, \qquad Q = \left\{\frac{5}{13}, \frac{8}{13}\right\}, \qquad v = \frac{20}{13}.$$

В следующем примере информационные множества выглядят немного иначе.

Пример 17.

1-й ход делает игрок A: он выбирает число x из множества двух чисел $\{1, 2\}$.

2-й ход делает игрок B: не зная о выборе игрока A на 1-м ходе, он выбирает число y из множества двух чисел $\{1, 2\}$.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, зная выбор у игрока B на 2-м ходе, но не помня собственного выбора x на 1-м ходе.

После этого игроки расплачиваются по правилу, указанному в примере 15.

Графическое представление этой игры показано на рис. 7.

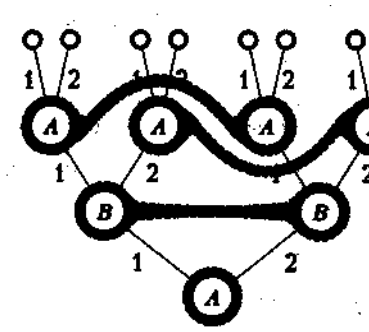


Рис. 7

Поскольку игроку B неизвестен выбор игрока A на 1-м ходе, то, выполняя свой ход, он не знает, в какой именно из двух возможных позиций он находится. Поэтому у игрока B всего две стратегии:

$$B_1 \text{ – выбрать } y = 1,\ B_2 \text{ – выбрать } y = 2$$

При описании стратегий игрока A нужно исходит из того, что к 3-му ходу игрок A утратил сведения о собственном выборе на 1-м ходе, но ему известен выбор игрока B на 2-м ходе. Поэтому выбор числа z игроку A следует связать с известным ему к 3-му ходу значением y. Удобнее всего это сделать по аналогии с расчетом стратегий игрока B в примерах 13 и 15, т. е. при помощи упорядоченной пары

$$[z_1, z_2]$$

Здесь z_1 ($z_1 = 1, 2$) — альтернатива, выбираемая игроком A при условии, что игрок B выбрал первую альтернативу, $y = 1$, а z_2 (z_2 — 1, 2) — альтернатива, выбираемая игроком A при условии, что игрок B выбрал вторую альтернативу, $y = 2$.

Чистую стратегию игрока A в данной игре можно записать так

$$(x, [z_1, z_2]).$$

Здесь x ($x = 1, 2$) — альтернатива, которую игрок A выбирает на 1-м ходе, z_1 ($z_1 = 1, 2$) — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал выбрал первую альтернативу ($y = 1$) и z_2 (z_2 — 1, 2) — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал вторую альтернативу ($y = 2$).

Например, выбор игроком A стратегии (2, [2, 1]) означает, что на 1-м ходе игрок A выбирает $x = 2$, а на 3-м $z = 2$, если игрок B выбрал $y = 1$, и z = 1, если игрок B выбрал $y = 2$.

Тем самым, у игрока A восемь чистых стратегий:

$$A_1 — (1, [1, 1]),\ A_2 — (1, [1, 2]), A_3 — (1, [2, 1]), A_4 — (1, [2, 2]),$$
$$A_5 — (2, [1, 1]),\ A_6 — (2, [1, 2]), A_7 — (2, [2, 1]), A_8 — (2, [2, 2]),$$

Покажем теперь, как в зависимости от применяемых стратегий определяются элементы таблицы выигрышей игрока A.

Пусть, например, игрок A выбрал стратегию A_3 — (1, [2, 1]), а игрок B — стратегию B_2 — (2). Тогда $x = 1$, $y = 2$, а из [2, 1] вытекает, что $z = 1$. Отсюда

$$W(x, y, z) = W(1, 2, 1) = 1.$$

По этой же схеме вычисляются и остальные элементы таблицы.

В результате получаем

		B_1	B_2
		(1)	(2)
A_1	(1, [1, 1])	W(1, 1, 1)	W(1, 2, 1)
A_2	(1, [1, 2])	W(1, 1, 1)	W(1, 2, 2)

A_3	(1, [2, 1])	W(1, 1, 2)	W(1, 2, 1)
A_4	(1, [2, 2])	W(1, 1, 2)	W(1, 2, 2)
A_5	(2, [1, 1])	W(2, 1, 1)	W(2, 2, 1)
A_6	(2, [1, 2])	W(2, 1, 1)	W(2, 2, 2)
A_7	(2, [2, 1])	W(2, 1, 2)	W(2, 2, 1)
A_8	(2, [2, 2])	W(2, 1, 2)	W(2, 2, 2)

$$\begin{pmatrix} -2 & 1 \\ -2 & -4 \\ 4 & 1 \\ 4 & -4 \\ 3 & -3 \\ 3 & 5 \\ 0 & -3 \\ 0 & 5 \end{pmatrix}$$

Оптимальные смешанны* стратегии игроков и цена игры соответственно равны:

$$P = \left\{0, 0, \frac{2}{5}, 0, 0, \frac{3}{5}, 0, 0\right\}, \qquad Q = \left\{\frac{4}{5}, \frac{1}{5}\right\}, \qquad v = \frac{17}{5}.$$

Рассмотрим позиционную игру со случайным ходом.

Пример 18.

1-й ход производится случайно: игрок O выбирает число x, равное 1, с вероятностью 0,5 и равное 2 с такой же вероятностью.

2-й ход делает игрок A: он выбирает число y из множества двух чисел {1, 2}, не зная результатов случайного выбора на 1-м ходе.

3-й ход делает игрок B: он выбирает число z из множества двух чисел {1, 2}, зная о том, какое именно число x случайно выбрано игроком O на 1-м ходе и не зная выбора y игрока A на 2-м ходе.

После этого игроки расплачиваются, используя функцию $W(x, y, z)$, ту же. что и в предыдущих примерах.

Графическое представление этой игры показано на рис.8.

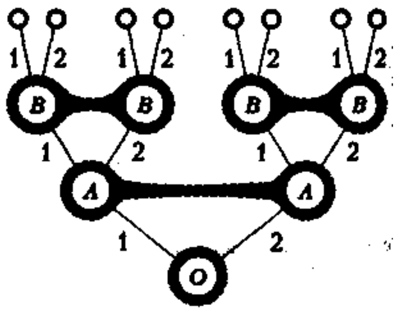


Рис. 8

Опишем стратегии игроков

Поскольку игроку A исход случайного испытания неизвестен, то он имеет всего две стратегии:

$$A_1 - (1),\ A_2 - (2),$$

При построении своих стратегий игроку B естественно воспользоваться имеющейся у него информацией о результате 1-го хода. Это позволит ему описать свою стратегию упорядоченной парой

$$[z_1, z_2].$$

Здесь z_1 ($z_1 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что $x = 1$, а z_2 ($z_2 = 1, 2$) — альтернатива, выбираемая игроком B при условии, что $x = 2$. Тем самым, у игрока B четыре стратегии:

$$B_1 — [1, 1],\ B_2 — [1, 2],\ B_3 — [2, 1],\ B_4 — [2, 2].$$

Покажем теперь, как определяются элементы таблицы выигрышей игрока A. Пусть, например, игрок A выбрал стратегию A_1 — (1), а игрок B — стратегию B_3 — [2, 1].

Различаются два случая

$$1)\ x = 1 \quad \text{и} \quad 2)\ x = 2.$$

Если $x = 1$, то стратегия B_3 указывает игроку B_3 его Выборг $z = 2$. А так как $y = 1$, то в результате имеем

$$W(x, y, z) = W(1, 1, 2) = 4.$$

Если $x = 2$, то стратегия B_3 указывает игроку B его выбор $z = 1$. А так как $y = 1$, то в результате

$$W(x, y, z) = W(2, 1, 1) = 3.$$

Поскольку первая и вторая альтернативы на 1-м ходе выбираются с вероятностями 0,5 и 0,5, то и вышеуказанные выигрыши появляются с теми же вероятностями и, следовательно, средний выигрыш игрока A при этих стратегиях определяется так

$$4 \text{ x } 0{,}5 + 3 \text{ x } 0{,}5 = 3{,}5.$$

Аналогичным образом рассчитывая остальные средние выигрыши, получаем при $x = 1$

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1)	$W(1, 1, 1)$	$W(1, 1, 1)$	$W(1, 1, 2)$	$W(1, 1, 2)$
A_2	(2)	$W(1, 2, 1)$	$W(1, 2, 1)$	$W(1, 2, 2)$	$W(1, 2, 2)$

или

$$\begin{pmatrix} -2 & -2 & 4 & 4 \\ 1 & 1 & -4 & -4 \end{pmatrix},$$

при $x = 2$

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1)	$W(2, 1, 1)$	$W(2, 1, 1)$	$W(2, 1, 2)$	$W(2, 1, 2)$
A_2	(2)	$W(2, 2, 1)$	$W(2, 2, 2)$	$W(2, 2, 2)$	$W(2, 1, 1)$

или

$$\begin{pmatrix} 3 & 0 & 3 & 0 \\ -3 & 5 & -3 & 5 \end{pmatrix}.$$

Искомая матрица игры имеет следующий вид

$$\begin{pmatrix} 0{,}5 & -1 & 3{,}5 & 0 \\ -1 & 3 & -3{,}5 & 0{,}5 \end{pmatrix}.$$

Наконец, рассмотрим пример позиционной игры со случайным разыгрыванием права первого хода.

Пример 19.

1-й ход делает игрок O, выбирая число x, равное 1 с вероятностью 2/3 и равное 2 с вероятностью 1/3.

Если x = 1, то на 2-м ходе игрок A выбирает число у из множества двух чисел {1, 2}, зная результат случайного выбора на 1-м ходе, а на **3-й ходе** игрок B выбирает число z из множества двух чисел {1, 2}, зная x, но не зная y.
Если x = 2, то на 2-м ходе игрок B выбирает число у из множества двух чисел {1, 2}, зная результат случайного выбора на 1-м ходе, а на **3-м ходе** игрок A выбирает число z из множества двух чисел {1, 2}, зная x, но не зная y.
После этого игроки расплачиваются, используя функцию $W(x, y, z)$, ту же, что и в предыдущих примерах.
Графическое представление этой игры показано на рис. 9.

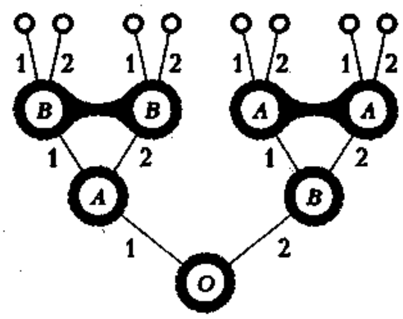


Рис. 9

Чистую стратегию игрока A в данной игре можно описать упорядоченной парой

$$|y, z|,$$

где y (y = 1, 2) — выбор игрока A на 2-м ходе, если на 1-м ходе выбрано x = 1, а z (z = 1, 2) — выбор игрока A на 3-м ходе, если на 1-м ходе выбрано x = 2.
Например, стратегия |1, 2| означает, что на 2-м хода игрок A выбирает y = 1, а на 3-м ходе — z = 2.

Тем самым, у игрока A четыре стратегии:

$$A_1 — |1, 1|, A_2 — |1, 2|, A_3 — |2, 1|, A_4 — |2, 2|.$$

У игрока B те же четыре стратегии:

$$B_1 — |1, 1|, B_2 — |1, 2|, B_3 — |2, 1|, B_4 — |2, 2|.$$

Покажем теперь, как находятся элементы матрицы выигрышей игрока A.
Пусть, например, игрок A применяет стратегию A_2 — |1, 2|, а игрок B — стратегию B_3 — |2, 1|.
Различаются два случая

1) $x = 1$ и 2) $x = 2$.

По условию при x = 1 игрок A имеет возможность сделать только 2-й ход (выбрать y), а игрок B — только 3-й (выбрать 2). При x = 2 их возможности меняются местами: игроку B предоставлено право 2-го хода (выбрать y), а игроку A — 3-го (выбрать z). .

Если x = 1, то стратегия A_2 указывает игроку A при 2-м ходе взять y = 1, а стратегия B_3 указывает игроку B при 3-м ходе взять z = 1. В результате

$$W(x, y, z) = W(1, 1, 1) = -2.$$

Если x = 2, то стратегия B_3 указывает игроку B при 2-м ходе взять y = 2, а стратегия A_2 указывает игроку A при 3-м ходе взять z = 2. В результате

$$W(x, y, z) = W(2, 2, 2) = 5.$$

Поскольку первая и вторая альтернативы на 1-м ходе выбираются соответственно с вероятностями 2/3 и 1/3, то и найденные выигрыши появляются с теми же вероятностями. Следовательно, математическое ожидание выигрыша игрока А при таких стратегиях рассчитывается так

$$-2\cdot\frac{2}{3}+5\cdot\frac{1}{3}=\frac{1}{3}\cdot$$

Итак,
при x = 1

		B_1	B_2	B_3	B_4
		|1, 1|	|1, 2|	|2, 1|	|2, 2|
A_1	|1, 1|	$W(1, 1, 1)$	$W(1, 1, 2)$	$W(1, 1, 1)$	$W(1, 1, 2)$
A_2	|1, 2|	$W(1, 1, 1)$	$W(1, 1, 2)$	$W(1, 1, 1)$	$W(1, 1, 2)$
A_3	|2, 1|	$W(1, 2, 1)$	$W(1, 2, 2)$	$W(1, 2, 1)$	$W(1, 2, 2)$
A_4	|2, 2|	$W(1, 2, 1)$	$W(1, 2, 2)$	$W(1, 2, 1)$	$W(1, 2, 2)$

или

$$\begin{pmatrix} -2 & 4 & -2 & 4 \\ -2 & 4 & -2 & 4 \\ 1 & -4 & 1 & -4 \\ 1 & -4 & 1 & -4 \end{pmatrix},$$

при $x = 2$

		B_1	B_2	B_3	B_4
		|1, 1|	|1, 2|	|2, 1|	|2, 2|
A_1	|1, 1|	$W(2, 1, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$	$W(2, 2, 1)$
A_2	|1, 2|	$W(2, 1, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$	$W(2, 2, 2)$
A_3	|2, 1|	$W(2, 1, 1)$	$W(2, 1, 1)$	$W(2, 2, 1)$	$W(2, 2, 1)$
A_4	|2, 2|	$W(2, 1, 2)$	$W(2, 1, 2)$	$W(2, 2, 2)$	$W(2, 2, 2)$

или

$$\begin{pmatrix} 3 & 3 & -3 & -3 \\ 0 & 0 & 5 & 5 \\ 3 & 3 & -3 & -3 \\ 0 & 0 & 5 & 5 \end{pmatrix}.$$

Отсюда получаем искомую матрицу игры

$$\frac{1}{3}\begin{pmatrix} -1 & 11 & -7 & 5 \\ -4 & 8 & 1 & 11 \\ 5 & -5 & -1 & -11 \\ 2 & -8 & 7 & -5 \end{pmatrix}.$$

Замечание. Графическое представление и функция выигрышей полностью определяют позиционную игру. В рассмотренных выше примерах 16-19 мы пользовались одной и той же функцией и одним и тем же деревом. Отличие было только в маркировке вершин дерева и информационных множествах. При построении последних необходимо соблюдать два правила:

1) в одно информационное множество могут входить позиции только одного игрока,

2) цепь, определяющая партию игры, может иметь с информационным множеством не более одной общей позиции.

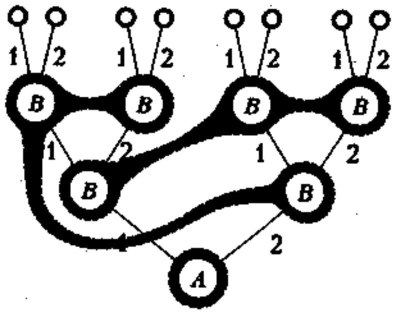


Рис. 10

Как показывает рис. 10, и при таких ограничениях информационные множества могут выглядеть довольно необычно.

3. Позиционные игры с полной информацией

Позиционная игра называется *игрой с полной информацией*, если в каждой позиции любой ее партии игрок, делающий ход, знает, какие альтернативы были выбраны на предыдущих ходах. В графическом описании каждая вершина дерева такой игры представляет собой отдельное информационное множество.

Примерами позиционных игр с полной информацией могут служить крестики-нолики, шашки и шахматы.

Основная особенность позиционной игры с полной информацией состоит в том, что соответствующая ей матрица выигрышей всегда имеет седловую точку, то есть в игре с полной информацией существуют оптимальные чистые стратегии и, значит, равновесная ситуация.

Сказанное означает, что в шахматах (крестиках-ноликах, шашках) уже в начальной позиции либо имеется способ выигрыша за белых, либо способ выигрыша за черных, либо как та, так и другая сторона способна форсировать ничью.

Однако известное доказательство существования равновесной ситуации неконструктивно и не дает эффективных приемов фактического нахождения решения игры.

И такие способы (стратегии) в шахматах не найдены до сих пор, и даже неизвестно, какая из перечисленных возможностей имеет место на самом деле.

Иное дело с игрой крестики-нолики: стратегий в ней немного и она разобрана до самого конца — существуют оптимальные чистые стратегии, ведущие игроков к ничьей.

Рассмотрим несколько **примеров**.

1. Как нетрудно заметить, двухходовая игра из примера 11 является игрой с полной информацией. Ее нормализация приводит к матрице с седловой точкой (см. пример 13).

2. **«Выкладывание монет на стол»**. Два игрока поочередно кладут монеты одинаковых размеров на обыкновенный стол, всякий раз выбирая произвольное доступное место для монеты (взаимное накрывание монет не допускается). Тот из игроков, кто положит монету, не оставляющую места для новых монет, выигрывает.

Это игра с полной информацией. Существует вполне определенная стратегия, обеспечивающая выигрыш тому из игроков, кто начинает игру. А именно, начинающий игру должен положить первую монету точно в центр стола и на каждый ход противника отвечать симметричным ходом. Исход игры от стратегии второго игрока не зависит.

3. **«Переговоры»**. В переговорах участвуют две стороны A и B. В слегка идеализированном варианте это может выглядеть, например, так.

Сначала сторона A высказывает одно из нескольких предложений, способных заинтересовать сторону B. Затем сторона B, ознакомившись с предложением стороны A, высказывает одно из нескольких встречных предложений, способных, по ее мнению, заинтересовать сторону A. В свою очередь, сторона A, ознакомившись с реакцией стороны B на сделанные предложения, высказывает ей новое предложение, внеся одну из нескольких возможных корректировок в свое первоначальное предложение с учетом мнения стороны B и т.д.

Если предмет переговоров сложен, то подобный обмен ходов может затянуться. Однако любые переговоры непременно заканчиваются. И там, на финише, ждет функция выигрышей.

Попробуем смоделировать короткий переговорный процесс трехходовой позиционной игрой.

Предположим, что переговоры заканчиваются через три хода, на каждом из которых соответствующая сторона имеет возможность выбора из двух альтернатив, и опишем соответствующую позиционную игру.

1-й ход делает сторона A: она выбирает одно из двух возможных предложений — число x из множества двух чисел $\{1, 2\}$.

2-й ход делает сторона B: она выбирает число у из множества двух чисел $\{1, 2\}$, зная число x, предложенное стороной A.

3-й ход делает сторона A: она выбирает число z из множества двух чисел $\{1, 2\}$, зная о предложении стороны B на 2-м ходе и помня собственное предложение на 1-м ходе.

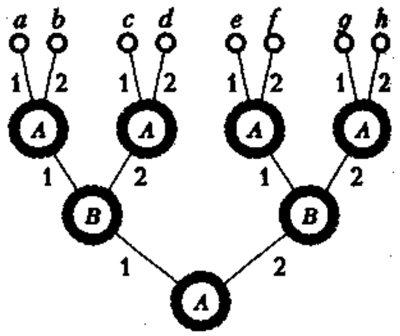


Рис. 11

После этого сторона A либо получает вознаграждение (например, в виде кредита от стороны B), либо выплачивает стороне B штраф.

Все эти возможности описываются функцией выигрышей $W(x, y, z)$, заданной следующей таблицей

$W(1, 1, 1) = a$	$W(2, 1, 1) = e$
$W(1, 1, 2) = b$	$W(2, 1, 2) = f$
$W(1, 2, 1) = c$	$W(2, 2, 1) = g$
$W(1, 2, 2) = d$	$W(2, 2, 2) = h$.

Графическое представление этой игры показано на рис. 11.

Ясно, что описанная позиционная игра является игрой с полной информацией. Начнем с описания возможных стратегий игрока В.

Поскольку игроку B выбор игрока A на 1-м ходе известен, то у игрока B те же четыре стратегии, что и в примере 13:

$$B_1 – [1, 1], B_2 – [1, 2], B_3 – [2, 1], B_4 – [2, 2],$$

С описания возможных стратегий игрока А дело обстоит немного посложнее — их восемь.

Чистая стратегия игрока A в данной игре описывается упорядоченной тройкой

$$(x, [z_1, z_2]).$$

Здесь x (x = 1, 2) — альтернатива, которую игрок A выбирает на 1-м ходе, z_1 (z_1 = 1, 2)— альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал первую альтернативу (y = 1) и z_2 (z_2 = 1, 2) — альтернатива, которую игрок A выбирает на 3-м ходе, если на 2-м ходе игрок B выбрал вторую альтернативу (y = 2).

Например, выбор игроком A стратегии (1, (2, 1]) означает, что на 1-м ходе игрок A выбирает $x = 1$, а на 3-м $z = 2$. если игрок B выбрал $y = 1$, и $z = 1$, если игрок B выбрал $y = 2$.

Тем самым, у игрока A восемь чистых стратегий:

A_1 — (1, [1, 1]), A_2 — (1, [1, 2]), A_3 — (1, [2, 1]), A_4 — (1, [2, 2]),

A_5 — (2, [1, 1]), A_6 — (2, [1, 2]), A_7 — (2, [2, 1]), A_8 — (2, [2, 2]),

Покажем теперь, как в зависимости от применяемых стратегий определяются элементы таблицы выигрышей игрока A.

Пусть, например, игрок A выбрал стратегию A_6 — (2, [1, 2]), а игрок B — стратегию B_3 – [2, 1]. Тогда $x = 2$. Из [2, 1] вытекает, что $y = 1$, а из (2, [1, 2]), что $z = 1$. Отсюда

$$W(x, y, z) = W(2, l, 1) = e.$$

Рассчитывая по этой же схеме все остальные элементы таблицы выигрышей, в итоге получим

		B_1	B_2	B_3	B_4
		[1, 1]	[1, 2]	[2, 1]	[2, 2]
A_1	(1, [1, 1])	W(1, 1, 1)	W(1, 1, 1)	W(1, 2, 1)	W(1, 2, 1)
A_2	(1, [1, 2])	W(1, 1, 1)	W(1, 1, 1)	W(1, 2, 2)	W(1, 2, 2)
A_3	(1, [2, 1])	W(1, 1, 2)	W(1, 1, 2)	W(1, 2, 1)	W(1, 2, 1)
A_4	(1, [2, 2])	W(1, 1, 2)	W(1, 1, 2)	W(1, 2, 2)	W(1, 2, 2)
A_5	(2, [1, 1])	W(2, 1, 1)	W(2, 2, 1)	W(2, 1, 1)	W(2, 2, 1)
A_6	(2, [1, 2])	W(2, 1, 1)	W(2, 2, 2)	W(2, 1, 1)	W(2, 2, 2)
A_7	(2, [2, 1])	W(2, 1, 2)	W(2, 2, 1)	W(2, 1, 2)	W(2, 2, 1)
A_8	(2, [2, 2])	W(2, 1, 2)	W(2, 2, 2)	W(2, 1, 2)	W(2, 2, 2)

$$\begin{pmatrix} a & a & c & c \\ a & a & d & d \\ b & b & c & c \\ b & b & d & d \\ e & g & e & g \\ e & h & e & h \\ f & g & f & g \\ f & h & f & h \end{pmatrix}$$

Вследствие того, что рассматриваемая позиционная игра является игрой с полной информацией, полученная матрица имеет седловую точку при любой функции выигрышей. В атом легко убедиться, произвольно выбирая значат» параметров *a, b, c, d, e, f, g и h.*

При увеличении числа ходов стратегии в позиционной игре с полной информацией строятся по аналогичной схеме.

Рассмотрим, например, четырехходовую позиционную игру.

1-й ход делает игрок A: он выбирает число x из множества двух чисел {1, 2}.

2-й ход делает игрок B: он выбирает число у из множества двух чисел {1, 2}, зная число x, выбранное игроком A на 1-м ходе.

3-й ход делает игрок A: он выбирает число z из множества двух чисел $\{1, 2\}$, зная число y, выбранное игроком B на 2-м ходе, и помня свой выбор числа x на 1-м ходе.
4-й ход делает игрок B: он выбирает число u из множества двух чисел $\{1, 2\}$, зная число z, выбранное игроком A на 3-м хода, помня свой выбор числа y на 2-м хода и зная выбор игрока A на 1-м ходе — число x.

После этого игроки A и B расплачиваются а соответствии с заданной функцией выигрышей $W(x, y, z, u)$.

В этой игре стратегии у игрока A те же, что и в задаче, рассмотренной выше: каждая из них задается тройкой вида

$$(x, [z_1, z_2]),$$

и общее их число равно восьми.

Что касается стратегий игрока B, то в этой игре их шестнадцать и каждая из них задается четверкой вида

$$([y_1, y_2], [u_1, u_2]).$$

Матрица выигрышей игрока A в данной игре имеет размер 8 х 16. Покажем, как определяются ее элементы в зависимости от применяемых стратегий игроков.

Пусть, например, игрок A выбрал стратегию A' — (2, [2, 1]), а игрок B стратегию B' — ([2, 1], [1,2]). Тогда $x = 1$, $y = 2$, а $z = 1$. Из того, что $[u_1, u_2] = [1, 2]$, получаем, что $u = 1$. Отсюда следует, что искомый элемент матрицы выплат равен

$$W(1, 2, 1, 1)$$

Остальные элементы матрицы вычисляются аналогично.

Так как эта позиционная игра также является игрой с полной информацией, то получаемая матрица будет иметь седловую точку.

Несколько слов в заключение

В рассмотренных примерах основное внимание было; уделено описанию процесса нормализации позиционной игры — построению дерева игры и информационных множеств, выработке стратегий игроков и вычислению элементов платежной матрицы. Следующий естественный шаг — отыскание цены игры и оптимальных стратегий игроков — проводится методами, о которых рассказывалось в главе, посвященной матричным играм.

Мы достаточно подробно остановились на позиционных играх двух лиц, где был: явно выражены интересы одного из игроков (игрока A). Следует, однако, иметь в вид] что в одних случаях интересы игрока B могут быть полностью противоположным интересам игрока A, в то время как в других вполне может оказаться, что то, что хорошо для одного игрока, не обязательно плохо для другого. Приведем два простых примера.

Пример А.
1-й ход. Игрок A выбирает число x из множества двух чисел $\{1, 2\}$.
2-й ход. Игрок B выбирает число y из множества двух чисел $\{1, 2\}$, зная выбор числа x игроком A.
Функции выплат игрокам A и B — $W_A(x, y)$ и $W_B(x, y)$ соответственно — задаются так:

$$W_A(1, 1) = 1, \quad W_A(1, 2) = -1, \quad W_A(2, 1) = -2, \quad W_A(2, 2) = 2,$$
$$W_B(1, 1) = 2, \quad W_B(1, 2)= 1, \quad W_B(2, 1) = 1, \quad W_B(2, 2) = 2.$$

Дерево игры показано на рис. 12. Исход игры зависит от того, каковы намерения игрока B — *максимизировать свой выигрыш*:

$$W_B(x, y) \rightarrow \max,$$

или *максимизировать свой относительный выигрыш*:

$$W_B(x, y)\ W_A(x, y)\text{- , y}) \rightarrow \max.$$

В первом Случав это достигается так:

При $x = 1$ $\quad y = 1$ и $W_B(1, 1) = 2$ $(W_A(1, 1) = 1)$;

При $x = 2$ $y = 2$ и W_B (2, 2) = 2 (W_A (2, 2) = 2).
Во втором случае:
При $x = 1$ $y = 2$ и W_B(1, 2) - W_A (1, 2) = 1 – (-1) = 2);
При $x = 2$ $y = 1$ и W_B (2, 1) - W_A (2, 1) = 1 – (-2) = 3).

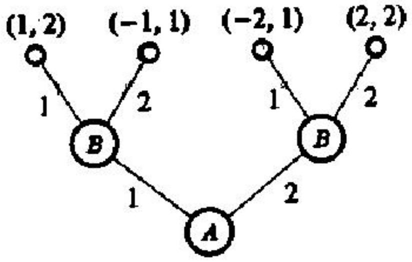


Рис.12

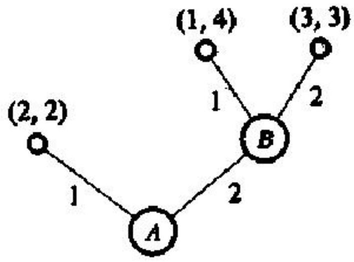


Рис. 13

Пример Б. Игра задается деревом (см. рис. 13).

1-й ход. Игрок A выбирает число x из множества двух чисел $\{1, 2\}$.

Если $x = 1$, то каждый из игроков получает свой выигрыш, равный 2.

Если $x = 2$, то право 2-го хода получает игрок B, где он и выбирает число y из множества двух чисел $\{1, 2\}$.

При $y = 1$ выигрыш игрока A равен 1, а игрока B — 4. При $y = 2$ оба игрока получают поровну — по 3.

В случае, когда каждый из игроков стремится к получению максимального выигрыша и любые виды кооперации запрещены, исход игры ясен — игрок A выбирает $x = 1$, и игра заканчивается. Но при x =2 и $y = 2$ каждый из игроков получает по 3 (такой исход предпочтительнее простейшего (1, 1)), и, если допустить соглашение между игроками, это обстоятельство вполне может изменить исход игры.

3.6 Принятие решений и теория игр. Примеры.

В теории игр рассматриваются ситуации, связанные с принятием решений, в которых два разумных противника имеют конфликтующие цели. К числу типичных примеров относится рекламирование конкурирующих товаров и планирование военных стратегий противоборствующих армий. Эти ситуации принятия решений отличаются от рассмотренных ранее, где природа не рассматривается в роли недоброжелателя.

В игровом конфликте участвуют два противника, именуемые **игроками**, каждый из которых имеет некоторое множество (конечное или бесконечное) возможных выборов, которые называются **стратегиями**. С каждой парой стратегий связан **платеж**, который один из игроков выплачивает другому. Такие игры известны как **игры двух лиц с нулевой суммой**, так как выигрыш одного игрока равен проигрышу другого. В такой игре достаточно задать результаты в виде платежей для одного из игроков. При обозначении игроков через A и B с числом стратегий m и n соответственно, игру обычно представляют в виде матрицы платежей игроку A:

	B_1	B_2	…	B_n
A_1	a_{11}	a_{12}	…	a_{1n}
A_2	a_{21}	a_{22}	…	a_{2n}
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
A_m	a_{m1}	a_{m2}	…	a_{mn}

Такое представление матричной игры означает, что если игрок A использует стратегию i, а игрок B — стратегию j, то платеж игроку A составляет a_{ij} и, следовательно, игроку B — $-a_{ij}$.

3.6.1. Оптимальное решение игры двух лиц с нулевой суммой

Поскольку игры берут свое начало в конфликте интересов, оптимальным решением игры является одна или несколько таких стратегий для каждого из игроков, при этом любое отклонение от данных стратегий не улучшает плату тому или другому игроку. Эти решения могут быть в виде единственной **чистой** стратегии или нескольких стратегий, которые являются **смешанными** в соответствии с заданными вероятностями. Рассматриваемые ниже примеры демонстрируют перечисленные случаи.

Пример 3.6-1

Две компании A и B продают два вида лекарств против гриппа; Компания A рекламирует продукцию на радио (A_1), телевидении (A_2) и в газетах (A_3) – Компания B, в дополнение к использованию радио (B_1), телевидения (B_2) и газет (B_3), рассылает также по почте брошюры (B_4). В зависимости от умения и интенсивности проведения рекламной кампании, каждая из компаний может привлечь на свою сторону часть клиентов конкурирующей компании. Приведенная ниже матрица характеризует процент клиентов, привлеченных или потерянных компанией A.

	B_1	B_2	B_3	B_4	Минимумы строк
A_1	8	–2	9	–3	–3
A_2	6	5	6	8	**5**← Максимин
A_3	–2	4	–9	5	–9
Максимумы столбцов	8	**5**	9	8	
		Минимакс			

Решение игры основано на обеспечении *наилучшего результата из наихудших* для каждого игрока. Если компания A выбирает стратегию A_1, то, независимо от того, что предпринимает компания B, наихудшим результатом является потеря компанией A 3% рынка в пользу компании B. Это определяется минимумом элементов первой строки матрицы платежей. Аналогично при выборе стратегии A_2 наихудшим исходом для компании A является увеличение рынка на 5% за счет компании B. Наконец, наихудшим исходом при выборе стратегии A_3 является потеря компанией A 9% рынка в пользу компании B. Эти результаты содержатся в столбце "Минимумы строк". Чтобы достичь наилучшего результата из наихудших, компания A выбирает стратегию A_2, так как она соответствует наибольшему элементу столбца "Минимумы строк".

Рассмотрим теперь стратегии компании B. Так как элементы матрицы являются платежами компании A, критерий *наилучшего результата из наихудших* для компании B соответствует выбору минимаксного значения. В результате приходим к выводу, что выбором компании B является стратегия B_2.

Оптимальным решением игры является выбор стратегий A_2 и B_2, т.е. обеим компаниям следует проводить рекламу на телевидении. При этом выигрыш будет в пользу компании A, так как ее рынок увеличится на 5%. В этом случае говорят, что **цена игры** равна 5% и что компании A и B используют стратегии, соответствующие **седловой точке**.

Решение, соответствующее седловой точке, гарантирует, что ни одной компании нет смысла пытаться выбрать другую стратегию. Действительно, если компания B переходит к другой стратегий (B_1, B_3 или B_4), то компания A может сохранить свой выбор стратегии A_2, что приведет к большей потере рынка компанией B (6% или 8%). По тем же причинам компании A нет резона использовать другую стратегию, ибо если она применит, например, стратегию A_3, то компания B может использовать свою стратегию B_3 и увеличить свой рынок на 9%. Аналогичные выводы имеют место, если компания A будет использовать стратегию A_1.

Оптимальное решение игры, соответствующее седловой точке, не обязательными должно характеризоваться чистыми стратегиями. Вместо этого оптимальное решение может требовать смешивания случайным образом двух или более стратегий, как это сделано в следующем примере.

Пример 3.6-2

Два игрока A и B играют в игру, основанную на подбрасывании монеты. Игроки одновременно и независимо друг от друга выбирают герб ($Г$) или решку ($Р$). Если результаты двух подбрасываний монеты совпадают (т.е. $ГГ$ или $РР$), то игрок A получает один доллар от игрока B. Иначе игрок A платит один доллар игроку B.

Следующая матрица платежей игроку A показывает величины минимальных элементов строк и максимальных элементов столбцов, соответствующих стратегиям обоих игроков.

Максиминная и минимаксная величины (цены) для этой игры равны – 1 доллар и 1 доллар соответственно. Так как эти величины не равны между собой, игра не имеет решения в чистых стратегиях; В частности, если игрок A использует стратегию $A_Г$, игрок B выберет стратегию $B_Р$, чтобы получить от игрока A один доллар. Если это случится, игрок A может перейти к стратегии $A_Р$, чтобы изменить исход игры и получить один доллар от игрока B. Постоянное искушение каждого игрока перейти к другой стратегии указывает на то, что решение в виде чистой стратегии неприемлемо. Вместо этого оба игрока должны использовать надлежащую случайную комбинацию своих стратегий. В рассматриваемом примере оптимальное значение цены игры находится где-то между максиминной и минимаксной ценами для этой игры:

максиминная (нижняя) цена $\leq$ цена игры $\leq$ минимаксная (верхняя) цена.

Следовательно, в данном случае цена игры должна лежать в интервале [–1,1], измеряемом в долларах.

Упражнения 3.6,а

1. Определите решение, определяемое седловой точкой, соответствующие чистые стратегии и цену игры для следующих игр, в которых платежи заданы для игрока A.

a)

	B_1	B_2	B_3	B_4
A_1	8	6	2	8
A_2	8	9	4	5
A_3	7	5	3	5

b)

	B_1	B_2	B_3	B_4
A_1	4	–4	–5	6
A_2	–3	–4	–9	–2
A_3	6	7	–8	–9
A_4	7	3	–9	5

2. В следующих играх заданы платежи игроку A. Укажите область значений для параметров p и q, при которых пара (2, 2) будет седловой точкой в каждой игре.

a)

B_1	B_2	B_3

A_1	1	q	6
A_2	p	5	10
A_3	6	2	3

b)

	B_1	B_2	B_3
A_1	2	4	5
A_2	10	7	q
A_3	4	p	6

3. Укажите область, которой принадлежит цена игры в каждом из следующих случаев, предполагая, что платежи заданы для игрока *A*.

a)

	B_1	B_2	B_3	B_4
A_1	1	9	6	0
A_2	2	3	8	4
A_3	–5	–2	10	–3
A_4	7	4	–2	–5

b)

	B_1	B_2	B_3	B_4
A_1	–1	9	6	8
A_2	–2	10	4	6
A_3	5	3	0	7
A_4	7	–2	8	4

c)

	B_1	B_2	B_3
A_1	3	6	1
A_2	5	2	3
A_3	4	2	–5

d)

	B_1	B_2	B_3	B_4
A_1	3	7	1	3
A_2	4	8	0	–6
A_3	6	–9	–2	4

4. Две фирмы производят два конкурирующих товара. Каждый товар в настоящее время контролирует 50% рынка. Улучшив качество товаров, обе фирмы собираются развернуть рекламные кампании. Если они не будут этого делать, то существующее состояние рынка не изменится. Однако если какая-либо фирма будет более активно рекламировать свои товары, то другая фирма потеряет соответствующий процент своих потребителей. Исследование рынка показывает, что 50% потенциальных потребителей получают информацию посредством телевидения 30% – через газеты и 20% – посредством радио.

a) Сформулируйте задачу в виде игры двух лиц с нулевой суммой и выберите подходящие средства рекламы для каждой фирмы.

b) Укажите интервал значений, которому принадлежит цена игры. Может ли каждая фирма действовать с единственной чистой стратегией?

5. Пусть a_{ij} – (i, j)-й элемент платежной матрицы с m стратегиями игрока A и n стратегиями игрока B. Элементы платежной матрицы представляют собой платежи игроку A. Докажите, что

$$\max_i \min_j a_{ij} \le \min_j \max_i a_{ij}.$$

3.6.2. Решение матричных игр в смешанных стратегиях

Решение матричных игр в смешанных стратегиях может быть найдено либо графически, либо методами линейного программирования. Графический метод применим для решения игр, в которых хоть один игрок имеет две чистые стратегии. Этот метод интересен в том плане, что графически объясняет понятие седловой точки. Методами линейного программирования может быть решена любая игра двух лиц с нулевой суммой.

ГРАФИЧЕСКОЕ РЕШЕНИЕ ИГР. Рассмотрим игру 2 х n, в которой игрок A имеет две стратегии.

Игра предполагает, что игрок A смешивает стратегии A_1 и A_2 с соответствующими вероятностями x_1 и $1 - x_1$, $0 \le x_1 \le 1$. Игрок B смешивает стратегии B_1, B_2,..., B_n с вероятностями y_1, y_2, ...,y_n, где $y_j \ge 0$,$j = 1, 2,...,$ и, $y_1 + y_2 + ... + y_n = 1$. В этом случае ожидаемый выигрыш игрока A, соответствующий j-й чистой стратегии игрока B, вычисляется в виде

Следовательно, игрок A ищет величину x_1, которая максимизирует минимум ожидаемых выигрышей

$$\max_{x_1} \min_j \left\{(a_{1j} - a_{2j})x_1 + a_{2j}\right\}.$$

Пример 3.6-3

Рассмотрим следующую игру 2x4, в которой платежи выплачиваются игроку A.

	B_1	B_2	B_3	B_4
A_1	2	2	3	–1
A_2	4	3	2	6

Игра не .имеет решения в чистых стратегиях, и, следовательно, стратегии должны быть смешанными. Ожидаемые выигрыши игрока A, соответствующие чистым стратегиям игрока B, приведены в следующей таблице.

Чистые стратегии игрока *B*	**Ожидаемые выигрыши игрока *A***
1	$-2x_1 + 4$
2	$-x_1 + 3$
3	$x_1 + 2$
4	$-7x_1 + 6$

На рис. 14.6 изображены четыре прямые линии, соответствующие чистым стратегиям игрока B. Чтобы определить *наилучший результат из наихудших,* построения *нижняя* огибающая четырех указанных прямых (изображенная на рисунке толстыми линейными сегментами), которая представляет минимальный (наихудший) выигрыш для игрока A независимо от того, что делает игрок B. Максимум (наилучшее) нижней огибающей соответствует максиминному решению в точке $x_1^* = 0.5$. Эта точка определяется пересечением прямых 3 и 4. Следовательно, оптимальным решением для игрока A является смешивание стратегий A_1 и A_2 вероятностями 0.5 и 0.5 соответственно. Соответствующая цена игры v определяется подстановкой x_1 = 0.5 уравнение либо прямой 3, либо 4, что приводит к следующему.

$$v = \begin{cases} \dfrac{1}{2} + 2 = \dfrac{5}{2}, & \text{из уравнения прямой 3,} \\ -7\left(\dfrac{1}{2}\right) + 6 = \dfrac{5}{2}, & \text{из уравнения прямой 4.} \end{cases}$$

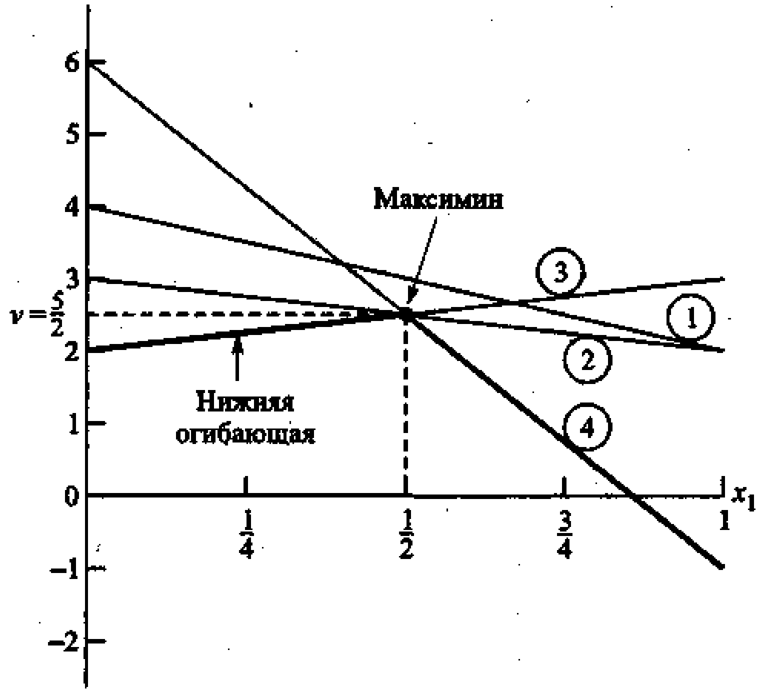


рис. 14.6

Оптимальная смешанная стратегия игрока *B* определяется двумя стратегиями, которые определяют нижнюю огибающую графика. Это значит, что игрок *B* может смешивать стратегии B_3 и B_4, в этом случае $y_1 = y_2 = 0$ и $y_4 = 1 - y_3$. Следовательно, ожидаемые платежи игрока *B,* соответствующие чистым стратегиям игрока *A,* имеют следующий вид.

Чистые стратегии игрока *B*	**Ожидаемые выигрыши игрока *A***
1	4ó$_3$ – 1
2	–4ó$_3$ + 6

Наилучшее решение из наихудших для игрока *B* представляет собой точку минимума *верхней* огибающей заданных двух прямых (построение прямых и определение верхней огибающей будет для Вас поучительным). Эта процедура эквивалентна решению уравнения

$$4y_3 - 1 = -4y_3 + 6.$$

Его решением будет y_3 = 7/8, что определяет цену игры v = 4 x (7/8) – 1 = 5/2.

Таким образом, решением игры для игрока *A* является смешивание стратегий A_1 и A_2 равными вероятностями 0.5 и 0.5, а для игрока *B* – смешивание стратегий B_3 и B_4 с вероятностями 7/8 и 1/8. (В действительности игра имеет альтернативное решение для игрока B, так как максиминная точка на рис. 14.6 определяется более чем двумя прямыми. Любая выпуклая линейная комбинация этих альтернативных решений также является решением задачи.)

Для игры, в которой игрок A имеет m стратегий, а игрок B – только две, решение находится аналогично. Главное отличие состоит в том, что здесь строятся графики функций, представляющих ожидаемые платежи второго игрока, соответствующие чистым стратегиям игрока A. В результате ведется поиск минимаксной точки *верхней огибающей* построенных прямых.

Упражнения 3.6,b

1. Решите графически игру с подбрасыванием монет из примера 3.6-2.

2. Робин часто путешествует между двумя городами. При этом есть возможность выбирать один из двух маршрутов: маршрут A представляет собой скоростное шоссе в четыре полосы, маршрут B – длинную обдуваемую ветром дорогу. Патрулирование дорог осуществляется ограниченным числом полицейских. Если все полицейские расположены на одном маршруте, Робин с ее страстным желанием ездить очень быстро, несомненно, получит штраф в 100 долларов за превышение скорости. Если полицейские патрулируют на двух маршрутах в соотношении 50 на 50, то имеется 50%-ная вероятность, что Робин получит штраф в 100 долларов на маршруте A и 30%-ная вероятность, что она получит такой же штраф на маршруте B. Кроме того, маршрут B длиннее, поэтому бензина расходуется на 15 долларов больше, чем на маршруте A. Определите стратегию как для Робин, так и для полиции.

3. Решите графически следующие игры, в которых платежи выплачиваются игроку A.

a)

	B_1	B_2	B_3
A_1	1	–3	7
A_2	2	4	–6

b)

	B_1	B_2
A_1	5	8
A_2	6	5
A_3	5	7

4. Дана следующая игра двух лиц с нулевой суммой.

	B_1	B_2	B_3
A_1	5.0	50.0	50.0
A_2	1.0	1.0	0.1
A_3	10.0	1.0	10.0

a) Проверьте, что смешанные стратегии с вероятностями (1/6, 0, 5/6) для игрока A и с вероятностями (49/54, 5/54, 0) для игрока B являются оптимальными, и определите цену игры.

b) Покажите, что цена игры равна

$$\sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij} x_i y_j \,.$$

РЕШЕНИЕ МАТРИЧНЫХ ИГР МЕТОДАМИ ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ. Теория игр находится в тесной связи с линейным программированием, так как любую конечную игру двух лиц с нулевой суммой можно представить в виде задачи линейного программирования и наоборот. Дж. Данциг [2] отмечает, что когда в 1947 году создатель теории игр Дж. фон Нейман впервые ознакомился с симплекс-методом, он сразу установил эта взаимосвязь и обратил особое внимание на концепцию двойственности в линейном программировании. Этот раздел иллюстрирует решение матричных игр методами линейного программирования.

Оптимальные значения вероятностей x_i, i = 1, 2,..., m, игрока A могут быть определены путем решения следующей максиминной задачи.

$$\max_{x_i}\left\{\min\left(\sum_{i=1}^{m}a_{i1}x_i,\sum_{i=1}^{m}a_{i2}x_i,\ldots,\sum_{i=1}^{m}a_{in}x_i\right)\right\},$$
$$x_1+x_2+\ldots+x_m=1,$$
$$x_i\geq 0, i=1,2,\ldots,m.$$

Чтобы сформулировать эту задачу в виде задачи линейного программирования, положим

$$v=\min\left(\sum_{i=1}^{m}a_{i1}x_i,\sum_{i=1}^{m}a_{i2}x_i,\ldots,\sum_{i=1}^{m}a_{in}x_i\right).$$

Отсюда вытекает, что

$$\sum_{i=1}^{m}a_{ij}x_i\geq v,\ j=1,2,\ldots,n\,.$$

Следовательно, задача игрока A может быть записана в виде

Максимизировать $z = v$

при ограничениях

$$\sum_{i=1}^{m}a_{ij}x_i\geq v,\ j=1,2,\ldots,n\,,$$
$$x_1+x_2+\ldots+x_m=1,$$
$$x_j\geq 0,\ j=1,2,\ldots,m,$$

v не ограничено в знаке.

Отметим последнее условие, что цена игры v может быть как положительной, так и отрицательной.

Оптимальные стратегии y_1, y_2, ...,y_n игрока B определяются путем решения задачи

$$\min_{y_j}\left\{\max\left(\sum_{j=1}^{n}a_{1j}y_j,\sum_{j=1}^{n}a_{2j}y_j,\ldots,\sum_{j=1}^{n}a_{mj}y_j,\right)\right\}$$

$$y_1+y_2+\ldots+y_n=1,$$
$$y_j\geq 0,\ j=1,2,\ldots,n.$$

Используя процедуру, аналогичную приведенной выше для игрока A, приходим к выводу, что задача для игрока B сводится к следующему.

Минимизировать $w = v$

при ограничениях

$$\sum_{j=1}^{m} a_{ij} y_j \le v, i = 1, 2, ..., m,$$

$$y_1 + y_2 + ... + y_n = 1,$$
$$y_j \ge 0,\ j = 1, 2, ..., n,$$

v не ограничена в знаке.

Две полученные задачи оптимизируют одну и ту же (не ограниченную в знаке) переменную v, которая является ценой игры. Причиной этого является то, что задача игрока B является двойственной к задаче игрока A (вам предлагается доказать это утверждение в упр. 3.5,с(6), используя определение двойственности). Это означает, что оптимальное решение одной из задач автоматически определяет оптимальное решение другой.

Пример 3.6-4

Решим следующую матричную игру методами линейного программирования.

	B_1	B_2	B_3	Минимумы строк
A_1	3	–1	–3	–3
A_2	–2	4	–1	**–2**
A_3	–5	–6	2	–6
Максимумы столбцов	3	4	**2**	

Значение цены игры v находится между –2 и 2.

Задача линейного программирования для игрока A

Максимизировать $z = v$

при ограничениях

$$3x_1 - 2x_2 - 5x_3 - v \ge 0,$$
$$-x_1 + 4x_2 - 6x_3 - v \ge 0,$$
$$-3x_1 - x_2 + 2x_3 - v \ge 0,$$
$$x_1 + x_2 + x_3 = 1,$$
$$x_1, x_2, x_3 \ge 0,$$

v не ограничено в знаке.

Оптимальным решением, полученным с помощью программы TORA, является $x_1 = 0.3945$, $x_2 = 0.3119$, $x_3 = 0.2936$ и $v = – 0.9083$.

Соответствующими двойственными переменными являются $y_1 = –0.3211$, $y_2 = –0.0826$, $y_3 = –0.5963$. Причина того, что переменные y_1, y_2, y_3 не являются положительными, как это должно быть, заключается в том, что задача линейного программирования для игрока A является задачей максимизации с ограничениями вида "≥". При этих условиях, как известно, соответствующе двойственные переменные должны быть отрицательными. Чтобы убедиться в том, что причина именно в этом, преобразуем все ограничения вид "≥" в задаче линейного программирования для игрока A в ограничения вида "≤" путем умножения каждого неравенства на –1. Соответствующие двойственные переменные будут неотрицательными, как и требуется (см. упр. 3.6,c(1)). Действительно, построение двойственной задачи непосредственно из задачи линейного программирования для игрока A показывает (см. упр. 3.6,c(6)), что в двойственной задаче, являющейся соответствующей задачей линейного программирования для игрока B, Должны быть $y_j \le 0$, но в то же время требуется выполнение условия $-y_1 - y_2 - ... - y_n = 1$, что равносильно требованиям $y_j \ge 0$ и $y_1 + y_2 + ... + y_n = 1$. К счастью, проблем, связанных со знаками, можно избежать преобразуя ограничения–неравенства вида "≥" в задаче линейного программирования для игрока A в ограничения–неравенства вида "≤".

Задача линейного программирования для игрока B

Минимизировать $z = v$

при ограничениях

$$3y_1 - y_2 - 3y_3 - v \le 0,$$
$$-2y_1 + 4y_2 - y_3 - v \le 0,$$
$$-5y_1 - 6y_2 + 2y_3 - v \le 0,$$
$$y_1 + y_2 + y_3 = 1,$$
$$y_1, y_2, y_3 \ge 0,$$

v не ограничено в знаке.

Оптимальным решением, полученным с помощью программы TORA, является $y_1 = 0.3211$, $y_2 = 0.0826$, $y_3 = 0.5963$ и $v = -0.9083$. Соответствующими двойственными переменными являются $x_1 = -0.3945$, $x_2 = -0.3119$, $x_3 = -0.2936$. Двойственные переменные отрицательны, ибо, подобно разобранной выше задаче для игрока A, в этом случае задача минимизации линейного программирования имеет ограничения–неравенства вида "≤". Приведение ограничений к виду "≥" исправит ситуацию со знаками.

Упражнений 3.6,c

1. Покажите в задаче из примера 3.6-4, что если ограничения–неравенства задачи линейного программирования для игрока A приведены к виду "≤", то аналогичные ограничения задачи для игрока B преобразуются в вид "≥" и двойственные переменные, полученные из одной или другой задачи, будут неотрицательными.

2. На загородном пикнике две команды, по два человека в каждой, играют в прятки. Есть четыре места, где можно спрятаться (А, Б, В и Г), и два члена прячущейся команды могут спрятаться каждый отдельно в любых двух из четырех мест. Затем другая команда имеет возможность проверить любые два места. Команда, которая ищет, получает премию, если будут обнаружены оба участника прячущейся команды, если же не обнаружен ни один участник, то она выплачивает премию. Иначе игра заканчивается вничью.

a) Сформулируйте задачу в виде игры двух лиц с нулевой суммой.
b) Определите оптимальные стратегии и цену игры.
c) Имеет ли задача альтернативные решения?
d) Является ли эта игра справедливой, т.е. имеет ли она цену, равную нулю?

3. Университетские команды UA и DU определяют свои стратегии игры в национальном чемпионате по баскетболу для колледжей. Оценивая возможности своих "запасных скамеек", каждый тренер разработал по четыре варианта замены игроков на протяжении игры. Способность каждой команды выполнять двух-, трех–очковые и штрафные броски является основным фактором, определяющим результат игры.

Приведенная ниже таблица содержит очки чистого выигрыша команды UA на протяжении одного владения мячом в зависимости от стратегий, планируемых каждой командой.

	DU_1	DU_2	DU_3	DU_4
UA_1	3	–2	1	2
UA_2	2	3	–3	0
UA_3	–1	2	–2	2
UA_4	–1	–2	4	1

a) Решите игру методами линейного программирования и определите выигрышные стратегии.

b) Исходя из имеющейся информации, какая из двух команд может выиграть чемпионат?

c) Пусть за всю игру имеется 60 возможностей владения мячом (30 владений для каждой команды). Предскажите ожидаемое количество очков, с которым будет выиграна игра чемпионата.

4. Армия полковника Блотто сражается с вражеской армией за контроль над двумя стратегически важными позициями. Полковник имеет в своем распоряжении два полка, а его противник — три. Каждый из противников может посыпать на любую позицию только целое число полков или совсем не посылать. Позиция будет захвачена армией, которая атакует большим количеством полков. Иначе результат сражения является ничейным

a) Сформулируйте задачу в виде игры двух лиц с нулевой суммой и решите игр; методами линейного программирования.

b) Какая армия выиграет сражение?

5. В игре двух лиц, именуемой двухпальцевой игрой Морра, каждый игрок показывает один или два пальца и одновременно отгадывает число пальцев, которые покажет его противник. Игрок, который угадал, выигрывает сумму, равную суммарному числу показанных противниками пальцев. Иначе игра заканчивается вничью. Сформулируйте задачу в виде игры двух лиц с нулевой суммой и решите игру методами линейного программирования.

6. Покажите, что задача, двойственная к задаче линейного программирования для игрока *A,* является задачей линейного программирования для игрока *B* и что еле дующие два утверждения не противоречат друг другу.

a) Задача линейного программирования для игрока *A* записана в форме, приведенной в разделе 3.6.2.

b) Задача линейного программирования для игрока *A* записана в форме, упомянутой в п. а), в которой все ограничения вида "≥" приведены к виду "≤".

3.7. Заключение

В этой главе рассмотрены критерии выбора оптимальных решений как в ситуациях точными (детерминированными), так и неполными (вероятностными или неопределенными) данными. Дальнейшие приложения критериев выбора оптимальных решений в ситуациях с неполными данными будут рассмотрены в последующих главах. В частности, критерий ожидаемой величины используется в главах, посвященных управлению запасам) теории массового обслуживания, имитационному моделированию и марковским процесса принятия решений. Это вовсе не значит, что другие критерии неприменимы. Критерий ожидаемой величины используется скорее традиционно в силу своей простоты.

Литература

Chen S. and Hwang C. *Fuzzy Multiple Decision Making*, Springer – Verlag, Berlin, 1992.

Dantzing G. B. *Linear Programming and Extension*, Princeton University Press, Princeton, N. J., 1963.

Meyerson R. *Game Theory: Analysis of Conflict*, Harvard University Press, Cambridge, Mass., 1991.

Saaty T. L. *Fundamentals of Decision Making*, RWS Publications, Pittsburg, 1994.

Комплексные задачи

1. Руководитель цеха рассматривает три возможных решения относительно существующего фрезерного станка.

a) Модифицировать имеющийся станок, установив на нем автоматическую подачу (АП).
b) Купить новый станок с программным управлением (ПУ).
c) Заменить станок обрабатывающим центром (ОЦ).

Три альтернативы оцениваются на основе двух критериев: денежный и функциональный. Следующая таблица содержит необходимые данные.

Критерий	**Единицы**	**АП**	**ПУ**	**ОЦ**
Денежный				
Начальная стоимость ($)		12 000	25 000	120 000
Стоимость обслуживания ($)		2 000	4 000	15 000
Стоимость обучения персонала ($)		3 000	8 000	20 000
Функциональный				
Производительность	Изделий/день	8	14	40
Время наладки	Минут	30	20	3
Металлические отходы	Фунты/день	440	165	44

Руководитель считает, что денежный критерий в полтора раза важнее функционального. Кроме того, производительность в два раза важнее времени наладки и в три раза важнее, чем количество получаемых металлических отходов. Показатель, связанный с временем наладки, считается в четыре раза важнее показателя, связанного с количеством металлических отходов. Что же касается денежного критерия, то руководитель считает, что стоимость обслуживания и стоимость обучения персонала одинаково важны, а начальная стоимость в два раза важнее каждого из этих двух показателей.

Проанализируйте описанную ситуацию и дайте соответствующие рекомендации.

2. Компания использует каталог товаров для продажи, включающий более 200 000 наименований, хранящихся на многих региональных складах. В прошлом компания считала важным иметь точный перечень запасов на каждом складе. Как следствие этого, каждый год проводился переучет — интенсивная и неприятная работа, которая неохотно выполнялась всеми складами. Компания для проверки качества складских операций в регионе сопровождала каждый переучет ревизией, которая охватывала около 100 наименований на каждом складе. Результаты проверки обнаружили, что в среднем лишь 64% наименований на каждом складе соответствовали действительной инвентарной описи, что

является неприемлемым. Дабы исправить ситуацию, компания распорядилась чаще проводить переучет дорогих и быстро реализуемых товаров. Системному аналитику была поставлена задача разработать процедуры для реализации этих планов. Вместо того чтобы напрямую заняться выполнением задания компании, системный аналитик решил установить причину возникшей проблемы. Он перешел в своем исследовании от формулировки "Как мы можем увеличить частоту переучетов?" к "Как можно повысить точность переучетов?". Изучение проблемы под таким углом зрения свелось к следующему анализу. Предполагая, что доля точно сосчитанных наименований на складе равна p, аналитик затем предположил, что есть основания считать, что имеется 95%-ная вероятность того, что если изделие было правильно учтено в первый раз, то будет правильно переучтено и при последующем переучете. Для части $1 - p$ товаров, которая не была точно учтена в первом раунде проверки, доля правильного учета во втором раунде равна 80%. Используя эту информацию, аналитик с помощью дерева решений построил график безубыточности, который сравнил точность учета в первом и втором раундах проверки. Конечный результат сводился к тому, что склады, на которых уровень точности выше порога безубыточности, не требовали переучета. Удивительным результатом предложенного решения было рьяное усилие со стороны каждого склада сделать правильный учет за первый раз, что привело к повышению точности учета на всех складах.

Как аналитик убедил руководство в жизнеспособности предложенного порога безубыточности для повторного переучета?

3. В аэрофлоте рабочие часы устанавливаются в соответствии с договорами, заключенными с профсоюзными организациями. В частности, максимальная продолжительность работы может быть ограничена 16 часами для полетов на Боинге-74719 (В-747) и 14 часами — на Боинге-707 (В-707). Если эти пределы превышаются в силу неожиданных задержек, экипаж должен быть заменен новым. Аэрофлот содержит резервные экипажи для таких случаев. Средняя годовая стоимость содержания член» резервного экипажа оценивается в 30 000 долларов. Задержка полета на одну ночь, обусловленная отсутствием резервного экипажа, может стоить 50 000 долларов. Член экипажа находится по вызову непрерывно 12 часов в сутки 4 дня в неделю и может находиться по вызову три оставшихся дня недели. Самолет В-747 может обслуживаться двумя экипажами для самолета В-707.

Следующая таблица содержит вероятности вызова резервных экипажей, вычисленные на основании трехлетнего опыта.

Категория рейса	Рейс (время вылета)	Вероятность вызова	
		В – 747	В – 707
1	14.0	0.014	0.072
2	13.0	0.000	0.019
3	12.5	0.000	0.006
4	12.0	0.016	0.006
5	11.5	0.003	0.003
6	11.0	0.002	0.003

Приведённые данные свидетельствуют, например, что для 14-часового рейса вероятность вызова равна 0.014 для В-747 и 0.072— для В-707.

Типичная пиковая часть расписания дня имеет следующий вид.

Время дня	**Самолет**	**Категория рейса**
8:00	707	3
9:00	707	6

	707	2
10:00	707	3
11:00	707	2
	707	4
15:00	747	6
16:00	747	4
19:00	747	1

Существующая политика относительно резервных экипажей состоит в использовании двух экипажей (по семь членов каждый) от 5:00 до 11:00, четырех — от 11:00 до 17:00 и двух — от 17:00 до 23:00.

Оцените эффективность Существующей политики относительно резервных экипажей. В частности, является ли число резервных экипажей очень большим, очень малым или таким, как необходимо?

С

Глава IV
ДЕТЕРМИНИРОВАННЫЕ МОДЕЛИ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

4.1. Введение

Динамическое программирование (ДП) определяет оптимальное решение n-мерной задачи путем ее декомпозиции на *n этапов,* каждый из которых представляет подзадачу относительно одной переменной. Вычислительное преимущество такого подхода состоит в том, что мы занимаемся решением одномерных оптимизационных подзадач вместо большой n-мерной задачи. Фундаментальным принципом ДП, составляющим основу декомпозиции задачи на этапы, является **оптимальность.** Так как природа каждого этапа решения зависит от конкретной оптимизационной задачи, ДП не предлагает вычислительных алгоритмов непосредственно для каждого этапа. Вычислительные аспекты решения оптимизационных подзадач на каждом этапе проектируются и реализуются по отдельности (что, конечно, не исключает применения единого алгоритма для всех этапов).

4.2 Рекуррентная природа вычислений ДП

Вычисления в ДП выполняются рекуррентно в том смысле, что оптимальное решение одной подзадачи используется в качестве исходных данных для следующей. Решив последнюю подзадачу, мы получим оптимальное решение исходной задачи. Способ выполнения рекуррентных вычислений зависит от того, как выполняется декомпозиция исходной задачи. В частности, подзадачи обычно связаны между собой некоторыми общими ограничениями. Если осуществляется переход от одной подзадачи к другой, то должны учитываться эти ограничения.

Пример 4.2-1. (Задача о кратчайшем пути)

Предположим, необходимо выбрать кратчайший путь между двумя городами. Сеть дорог, показанная на рис. 10.1, представляет возможные маршруты между исходным городом, находящимся в узле 1, и конечным пунктом, который находится в узле 7. Маршруты проходят через промежуточные города, обозначенные на сети узлами с номерами 2-6.

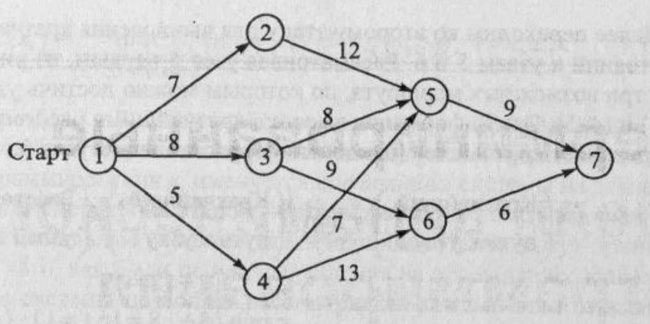
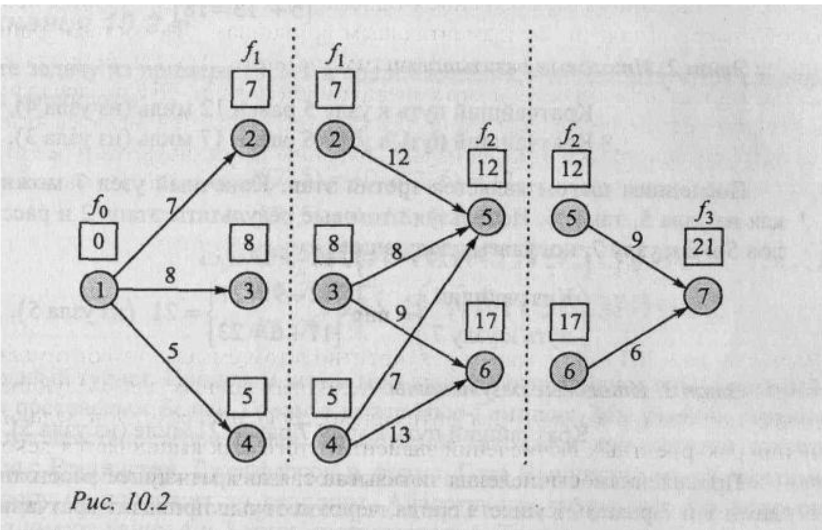


Рис. 10.1

Мы можем решить эту задачу посредством полного перебора всех маршрутов между узлами 1 и 7 (имеется пять таких маршрутов). Однако в большой сети полный перебор является неэффективным с вычислительной точки зрения.

Чтобы решить эту задачу с помощью методов динамического программирования, сначала разделим ее на *этапы.* Вертикальные пунктирные линии на рис. 10.2 очерчивают три этапа задачи. Далее выполняются вычисления для каждого этапа в отдельности.



Общая задача состоит в вычислении кратчайших (постепенно накапливаемых) расстояний ко всем вершинам этапа с последующим использованием этих расстояний в качестве исходных данных для следующего этапа. Рассматривая узлы, относящиеся к первому этапу, замечаем, что каждый из узлов 2, 3 и 4 связан с начальным узлом 1 единственной дугой (рис. 4.2). Следовательно, для первого этапа имеем следующее.

Этап 1. Итоговые результаты.

Кратчайший путь к узлу 2 равен 7 миль (из узла 1),

Кратчайший путь к узлу 3 равен 8 миль (из узла 1),

Кратчайший путь к узлу 4 равен 5 миль (из узла 1).

Далее переходим ко второму этапу для вычисления кратчайших (накопленных) расстояний к узлам 5 и 6. Рассматривая узел 5 первым, из рис. 10.2 замечаем, что есть три возможных маршрута, по которым можно достичь узла 5, а именно (2, 5), [(3, 5) и (4, 5). Эта информация вместе с кратчайшими расстояниями к узлам 2, 3, и 4 определяет кратчайшее (накопленное) расстояние к узлу 5 следующим образом.

$$\begin{pmatrix} \textit{Кратчайший} \\ \textit{путь к узлу } 5 \end{pmatrix} =$$

$$\min_{i=2,3,4}\left\{\begin{pmatrix} \textit{кратчайший} \\ \textit{путь к узлу } i \end{pmatrix} + \begin{pmatrix} \textit{расстояние от} \\ \textit{узла } i \textit{ к узлу } 5 \end{pmatrix}\right\} = \min\begin{Bmatrix} 7+12=19 \\ 8+8\ \ =16 \\ 5+7\ \ =12 \end{Bmatrix} = 12\ \ (\textit{из узла} 4)$$

Аналогично для узла 6 имеем следующее.

$$\begin{pmatrix} \textit{Кратчайший} \\ \textit{путь к узлу } 6 \end{pmatrix} =$$

$$\min_{i=3,4}\left\{\begin{pmatrix} \textit{кратчайший} \\ \textit{путь к узлу } i \end{pmatrix} + \begin{pmatrix} \textit{расстояние от} \\ \textit{узла } i \textit{ к узлу } 6 \end{pmatrix}\right\} = \min\begin{Bmatrix} 8+9\ \ =17 \\ 5+13=18 \end{Bmatrix} = 17\ \ (\textit{из узла } 3)$$

Этап 2. Итоговые результаты.

Кратчайший путь к узлу 5 равен 12 миль (из узла 4),

Кратчайший путь к узлу 6 равен 17 миль (из узла 3).

Последним шагом является третий этап. Конечный узел 7 можно достигнуть как из узла 5, так и 6. Используя итоговые результаты этапа 2 и расстояния от узлов 5 и 6 к узлу 7, получаем следующее:

$$\begin{pmatrix} \textit{Кратчайший} \\ \textit{путь к узлу } 7 \end{pmatrix} = \min\begin{Bmatrix} 12+9=21 \\ 17+6=23 \end{Bmatrix} = 21\ \ (\textit{из узла } 5)$$

Этап 3. Итоговые результаты.

Кратчайший путь к узлу 7 равен 21 миле (из узла 5).

Привеленные вычисления показывают, что кратчайшее расстояние между узлами 1 и 7 равно 211 миле. Города, через которые проходит кратчайший маршрут, определяется следующим образом. Из готовых результатов третьего этапа следует, что узел 7 связывается с узлом 5. Далее из итоговых результатов второго этапа следует, что узел 4 связывается с узлом 5. Наконец, из готовых результатов первого этапа следует, что узел 4 связывается с узлом 1. Следовательно, оптимальным маршрутом является последовательность 1-4-5-7.

Теперь покажем, как рекуррентные вычисления динамического программирования можно выразить математически. Пусть $f_i(x_i)$ – кратчайшее расстояние до вершины x_i на этапе $i, d(x_{i-1}, x_i)$ – расстояние от узла x_{i-1} до узла x_i. Тогда f_i вычисляется из f_{i-1} с помощью следующего рекуррентного уравнения.

$$f_i(x_i) \min_{\substack{\text{все допустимые} \\ (x_{i-1}, x_i)\text{маршруты}}} \left\{ d(x_{i-1}, x_i) + f_{i-1}(x_{i-1}) \right\}, i = 1, 2, 3.$$

При $i=1$ полагаем $f_0(x_0) \equiv 0$. Это уравнение показывает, что кратчайшие расстояния $f_i(x_i)$ на этапе i должны быть выражены как функции следующего узла x_i ,. В терминологии динамического программирования *x,* именуется *состоянием* системы на этапе i. В действительности *состояние* системы на этапе i — это информация, связывающая этапы между собой, при этом оптимальные решения для оставшихся этапов могут приниматься без повторной проверки того, как были получены решения на предыдущих этапах. Такое определение *состояния* системы позволяет рассматривать каждый этап отдельно и гарантирует, что решение является допустимым на каждом этапе.

Определение *состояния* системы приводит к следующему унифицированному положению.

Принцип оптимальности. На каждом этапе оптимальная стратегия определяется независимо от стратегий, примененных на предыдущих этапах.

Применение принципа оптимальности демонстрируется вычислениями из примера 10.2-1. Например, на этапе 3 мы используем кратчайшие пути к узлам 5 и 6 и не интересуемся, как эти узлы были достигнуты из узла 1.

Упражнения 4.2,а

Решите задачу из примера 10.2-1 в предположении, что используются следующие длины маршрутов:

$d(1,2) = 5$, $d(1,3) = 9$, $d(1,4) = 8$,
$d(2,5) = 10$, $d(1,6) = 17$,
$d(3,5) = 4$, $d(4,6) = 9$,
$d(5,7) = 8$,
$d(6,7) = 9$.

2.Я — заядлый турист. Прошлым летом мой друг и я отправились в пятидневный поход по прекрасным Белым Горам в штате Нью-Гемпшир. Мы решили ограничить наше путешествие территорией, на которой находится три хорошо известные вершины: Вашингтон, Джефферсон и Адаме. Гора Вашингтон имеет шестимильную тропу от подножия до вершины. Аналогичные тропы гор Джефферсона и Адамса имеют длину 4 и 5 миль соответственно. Тропы, соединяющие подножия этих трех гор, имеют следующую длину: 3 мили между вершинами Вашингтона и Джефферсона, 2 мили между вершинами Джефферсона и Адамса и 5 миль между вершинами Адамса и Вашингтона. В первый день мы стартовали от подножия вершины Вашингтона и вернулись в эту же точку к концу пятого дня. Нашей целью было пройти как можно более длинный путь. Мы также решили подниматься каждый день только на одну вершину и располагаться лагерем у подножия той горы, на которую мы решили восходить на следующий день. Кроме того, мы решили, что не будем подниматься на одну и ту же вершину в течение двух дней подряд. Каким было расписание нашего похода?

4.3. Рекуррентные алгоритмы прямой и обратной прогонки

В примере 4.2-1 вычисления проводились последовательно: от первого этапа до третьего. Такая последовательность вычислений известна как алгоритм прямой прогонки. Этот же пример может быть решен с помощью алгоритма обратной прогонки, в соответствии с которым вычисления проводятся от третьего этапа до первого.

Алгоритмы прямой и обратной прогонки приводят к одному и тому же решению. Несмотря на то, что алгоритм прямой прогонки представляется более логичным, в специальной литературе, посвященной динамическому программированию, неизменно используется алгоритм обратной прогонки. Причина этого в том, что в общем случае алгоритм обратной прогонки может быть более эффективным с вычислительной точки зрения. Продемонстрируем использование алгоритма обратной прогонки на примере 4.2-1. Мы также представим вычисления динамического программирования в компактной табличной форме.

Пример 3-1

Рекуррентное уравнение для алгоритма обратной прогонки в примере 4.2-1 имеет вид

$$f_i(x_i) \min_{\substack{\text{все допустимые} \\ (x_{i-1},x_i)\,\text{маршруты}}} \left\{ d(x_i, x_{i+1}) + f_{i+1}(x_{i+1}) \right\}, i = 1, 2, 3.$$

где $f_4(x_4) \equiv 0$ для $x_4 = 7$. Соответствующей последовательностью вычислений будет $f_3 \to f_2 \to f_1$.

Этап 3. Так как узел 7 (x_4 = 7) связан с узлами 5 и 6 *(*x_3 = *5* и 6) в точности одним маршрутом, альтернативы для выбора отсутствуют, а результаты третьего этапа можно подытожить следующим образом.

x_3	$\overline{d(x_3, x_4)}$	Оптимальное решение	
	$x_4 = 7$	$f_3(x_3)$	x_4
	9	9	7
	6	6	7

Этап 2. Так как маршрута (2,6) не существует, соответствующая альтернатива не рассматривается. Используя значения $f_3(x_3)$, полученные на третьем этапе, мы можем сравнить допустимые альтернативные решения, как показано в следующей таблице.

x_2	$\overline{d(x_2, x_3)} + f_3(x_3)$		Оптимальное решение	
	$x_3 = 5$	$x_3 = 6$	$f_2(x_2)$	x_3
	12+9=21	–	21	5
	8+9=17	9+6=15	15	6
	7+9=16	13+6=19	16	5

Оптимальное решение второго этапа означает следующее. Если вы находитесь в узле (городе) 2 или 4, кратчайший путь к узлу 7 проходит через узел 5, а если находитесь в узле 3, кратчайший путь к узлу 7 проходит через узел 6.

Этап I. Из узла 1 мы имеем три альтернативных маршрута: (1,2), (1,3) и (1,4). Используя значения $f_2(x_2)$, полученные на втором этапе, вычисляем данные следующей таблицы.

	$\overline{d(x_1,x_2)} + f_2(x_2)$			Оптимальное решение	
x_2	$x_2=2$	$x_2=3$	$x_2=4$	$f_1(x_1)$	x_2*
1	7+21=28	8+15=23	5+16=21	21	4

Оптимальное решение на первом этапе показывает, что кратчайший путь проходит через город 4. Далее из оптимального решения на втором этапе следует, что из города 4 необходимо двигаться в город 5. Наконец, из оптимального решения на третьем этапе следует, что город 5 связан с городом 7. Следовательно, полным маршрутом, имеющим кратчайшую длину, является 1-4-5-7, и его длина равна 21 миле.

Упражнения 4. 3,а

1. Для задачи из упр. 4.2,а(1) получите рекуррентное соотношение обратной прогонки и используйте его для получения оптимального решения.

2. Для задачи из упр. 4.2,а(2) получите рекуррентное соотношение обратной прогонки и используйте его для получения оптимального решения.

3. Определите кратчайший маршрут между городами 1 и 7 на сети дорог, представленной на рис. 10.3. Определите этапы и состояния системы с помощью алгоритма обратной прогонки, а затем решите задачу.

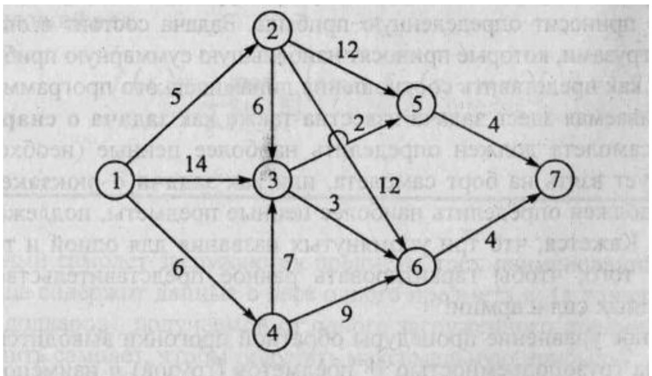


Рис. 10.3

4.4. Некоторые приложения динамического программирования

В данном разделе рассмотрено четыре примера, каждый из которых выбран для демонстрации методов динамического программирования. При рассмотрении каждого примера особое внимание обратите на три основных элемента моделей динамического программирования. 1. Определение *этапов.*

2. Определение на каждом этапе *вариантов решения* (альтернатив).

3. Определение *состояний* на каждом этапе.

Из перечисленных выше элементов снятие *состояния,* как правило, представляется весьма сложным для восприятия. Рассмотренные в этом разделе приложения последовательно показывают, что определение состояния меняется в зависимости от моделируемой ситуации. При рассмотрении каждого приложения полезно ответить на следующие вопросы.

1. Какие соотношения связывают этапы вместе?

2. Какая информация необходима для того, чтобы получить допустимые решения на текущем этапе без повторной проверки решений, принятых на предыдущих этапах?

Мой опыт преподавания показывает, что понятие *состояния* удается глубже уяснить, если поставить под сомнение определение состояния, которое предложено в настоящей книге. Рекомендуем воспользоваться каким-либо другим определением, которое покажется вам "более логичным", и применить его в рекуррентных вычислениях. В конечном счете вы сможете убедиться, что приведенные здесь определения обеспечивают корректное решение задач. В то же время преложенный подход будет способствовать вашему пониманию самой концепции состояния.

4.4.1. Задача о загрузке

Задача о загрузке — это задача о рациональной загрузке судна (самолета, автомашины и т.п.), которое имеет ограничения по объему или грузоподъемности. Каждый помещенный на судно груз приносит определенную прибыль. Задача состоит в определении загрузки судна такими грузами, которые принося наибольшую суммарную прибыль.

Перед тем как представить соотношения динамического программирования, заметим, что рассматриваемая здесь задача известна также как задача о снаряжении, где пилот реактивного самолета должен определить наиболее ценные (необходимые) предметы, которые следует взять на борт самолет», или как задача о рюкзаке, в которой солдат (или турист) должен определить наиболее ценные предметы, подлежащие загрузке в ранец (рюкзак). Кажется, что три упомянутых названия для одной и той же задачи были выбраны для того, чтобы гарантировать равное представительство военно-морского флота, воздушных сил и армии.

Рекуррентное уравнение процедуры обратной прогонки выводится для общей задачи загрузки судна грузоподъемностью W предметов (грузов) n наименований. Пусть m_i — количество предметов *i-го* наименования, подлежащих загрузке, r,- — прибыль, которую приносит один загруженный предмет *i-го* наименования, w_i — вес одного предмета *i-го* наименования. Общая задача имеет вид. следующей целочисленной задачи линейного программирования.

Максимизировать $z = r_1m_1 + r_2m_2 + ... + r_nm_n$

при условии, что

$$w_1m_1 + w_2m_2 + ... + w_nm_n \le W,$$

$$m_1, m_2, ..., m_n \ge 0 \text{ и целые.}$$

Три элемента модели динамического программирования определяются следующим образом.

1. *Этап i* ставится в соответствие предмету *i-го* наименования, $i = 1, 2, ..., n.$

2. *Варианты решения* на этапе описываются количеством m_i предметов *i-го* наименования, подлежащих загрузке. Соответствующая прибыль равна r_im_i. Значение m_i заключено в пределах от 0 до $[W / w_i]$, где $[W / w_i]$ — целая часть числа W / w_i

3. *Состояние* x_i на этапе i выражает суммарный вес предметов, решения о погрузке которых приняты на этапах $i, i+1, ..., n$. Это определение отражает тот факт, что ограничение по весу является единственным, которое связывает n этапов вместе.

Пусть $f_i(x_i)$ — максимальная суммарная прибыль от этапов $i, i+1, ..., n$ при заданном состоянии x_i. Проще всего рекуррентное уравнение определяется с помощью следующей двухшаговой процедуры.

Шаг 1. Выразим $f_i(x_i)$ как функцию $f_{i+1}(x_{i+1})$ в виде

$$f_i(x_i) = \min_{\substack{m_i = 0,1,...,\left[\frac{W}{w_i}\right] \\ x_i = 0,1,...,W}} \{r_i m_i + f_{i+1}(x_{i+1})\}, i = 1, 2, ..., n.$$

где $f_{n+1}(x_{n+1}) = 0.$

Шаг 2. Выразим x_{i+1} как функцию x_i,• для гарантии того, что левая часть последнего уравнения является функцией лишь x_i. По определению x_i - x_{i+1} представляет собой вес, загруженный на этапе i т.е. $x_i - x_{i+1} = w_i m_i$ или $x_{i+1} = x_i - w_i m_i$ Следовательно, рекуррентное уравнение приобретает следующий вид.

Пример 4.4-1

В 4-тонный самолет загружаются предметы трех наименований. Приведенная ниже таблица содержит данные о весе одного предмета w_i (в тоннах) и прибыли r_i (в тысячах долларов), получаемой от одного загруженного предмета. Как необходимо загрузить самолет, чтобы получить максимальную прибыль?

Предмет i	w_i	r_i
1	2	31
2	3	47
3	1	14

Так как вес одного предмета w_i для всех наименований и максимальный вес W принимают целочисленные значения, состояние x_i может принимать лишь целочисленные значения.

Этап 3. Точный вес, который может быть загружен на этапе 3 (предмет наименования 3), заранее неизвестен, но он должен принимать одно из значений 0,1,...,4 (так как W = 4 тонны). Состояния x_3 = 0 и x_3 = 4 представляют собой крайние случаи, когда предметы третьего наименования совсем не загружаются или загружают самолет полностью. Остальные значения x_3 (= 1, 2 или 3) предполагают частичную загрузку самолета предметами третьего наименования. Действительно, при этих значениях x_3 все оставшиеся емкости самолета могут быть заполнены предметами третьего наименования.

Так как вес w_3 одного предмета третьего типа равен 1 тонне, максимальное количество единиц этого типа, которое может быть загружено, равно [4/1] = 4. Это означает, что возможными значениями x_3 будут О, 1, 2, 3 и 4. Решение m_3 является допустимым лишь при условии, что $w_3 m_3 \le x_3$.Следовательно, все недопустимые альтернативы (те, для которых $w_3 m_3 > x_3$) исключены. Следующее уравнение является основой для сравнения альтернатив на этапе 3.

$$f_3(x_3) = \max_{m_3}\{14m_3\}, \max\{m_3\} = \left[\frac{4}{1}\right] = 4.$$

В следующей таблице сравниваются допустимые решения для каждого значения x_3.

x_3.	14 m_3					Оптимальное решение	
	m_3=0	m_3=1	m_3=2	m_3=3	m_3=4	$f_3(x_3)$	m_3
0	0	–	–	–	–	0	0

1	0	14	–	–	–	14	1
2	0	14	28	–	–	28	2
3	0	14	28	42	–	42	3
4	0	14	28	42	56	56	4

Этап 2.

$$f_2(x_2) = \max_{m_2}\left\{47m_2 + f_3(x_2 - 3m_2)\right\}, \max\left\{m_2\right\} = \left[\frac{4}{3}\right] = 1.$$

x_2	$47m_2 + f_3(x_2 - 3m_2)$		Оптимальное решение	
	$m_2 = 0$	$m_2 = 1$	$f_2(x_2)$	m_2
0	0+0=0	–	0	0
1	0+14=14	–	14	0
2	0+28=28	–	28	0
3	0+42=42	47+0=47	47	1
4	0+56=56	47+14=61	61	1

Этап 1.

$$f_1(x_1) = \max_{m_1}\left\{31m_1 + f_2(x_1 - 2m_1\right\}, \max\left\{m_1\right\} = \left[\frac{4}{2}\right] = 2$$

x_1	$31m_1 + f_2(x_1 + f_2(x_1 - 2m_1)$			Оптимальное решение	
	$m_1 = 0$	$m_1 = 1$	$m_1 = 2$	$f_1(x_1)$	m_1 *
0	0+0=0	–	–	0	0
1	0+14=14	–	–	14	0
2	0+28=28	31+0=31	–	31	1
3	0+47=47	31+14=45	–	47	0
4	0+61=61	31+28=59	62+0=62	62	2

Оптимальное решение определяется теперь следующим образом. Из условия W = 4 следует, что первый этап решения задачи при x_1 = *4* дает оптимальное решение m_1* - 2, которое означает, что два предмета первого наименования будут загружены в самолет. Эта загрузка оставляет $x_2 = x_1 - 2m_1* = 4 - 2 \times 2 = 0.$ Решение на втором этапе при x_2= 0 приводит к оптимальному решению m_2* - *0*, которое, в свою очередь, дает $x_3 = x_2 - 3m_2 = 0 - 3 \times 0 = 0$. Далее этап 3 при $x_3 = 0$ приводит к $m_3* = 0$. Следовательно, оптимальным решением задачи является $m_1* = 2$ $m_2* = 0$ $m_3* = 0$.Соответствующая прибыль равна 62 000 долларов.

В таблице для первого этапа нам, по существу, необходимо получить оптимальное решение лишь для $x_1 = 4$, так как это последний этап, подлежащий рассмотрению. Однако в таблицу включены также вычисления для $x_1 = 0,1,2$ и 3, которые позволяют провести анализ чувствительности решения. Например, что произойдет, если максимальная грузоподъемность самолета будет 3 тонны вместо 4? Новое оптимальное решение может быть определено, начиная с $x_1 = 3$ на первом этапе и продолжая так, как мы поступали при $x_1 = 4$.

Задача о загрузке является типичным представителем задачи *распределения ресурсов,* в которой ограниченный ресурс распределяется между конечным числом видов (экономической) деятельности. При этом целью является максимизация соответствующей функции прибыли. В таких моделях определение состояния на каждом этапе будет аналогично приведенному для задачи о загрузке: состоянием на этапе i является суммарное количество ресурса, распределяемого на этапах $i, i+1, ..., n.$

Упражнения 4. 4,а

1. В задаче примера 10.4-1 определите оптимальное решение, предполагая, что максимальная грузоподъемность самолета составляет 3 тонны.

2. Решите задачу о загрузке из примера 10.4-1 для каждого из следующих случаев.

- $w_1 = 4, r_1 = 70, w_2 = 1, r_2 = 20, w_3 = 2, r_3 = 40, W = 6.$
- $w_1 = 1, r_1 = 30, w_2 = 2, r_2 = 60, w_3 = 3, r_3 = 80, W = 4.$

3. Турист собирается в путешествие по дикой местности и должен упаковать в рюкзак предметы трех наименований: пищу, средства первой помощи и одежду. Объем рюкзака составляет 3 кубических фута. Каждая единица пищи занимает 1 кубический фут, упаковка средств первой помощи — четверть кубического фута, а отдельный предмет одежды — примерно половину кубического фута. Турист определил свои предпочтения весовыми коэффициентами 3, 4 и 5 — для пищи, средств первой помощи и одежды соответственно. Это означает, что одежда является самым ценным предметом среди остальных. Опыт подсказывает туристу, что он должен взять не менее одного предмета каждого наименования и не более двух комплектов средств первой помощи. Сколько единиц каждого наименования возьмет турист в поход?

4. Студент должен выбрать 10 факультативных курсов на четырех различных факультетах, причем на каждом факультете должен быть выбран по меньшей мере один курс. Эти курсы распределяются между факультетами таким образом, чтобы максимизировать объем "знаний". Студент оценивает знания по шкале в сто баллов и приходит к выводам, представленным в следующей таблице.

	Номер курса						
Факультет	1	2	3	4	5	6	≥ 7
I	25	50	60	80	100	100	100
II	20	70	90	100	100	100	100
III	40	60	80	100	100	100	100
IV	10	20	30	40	50	60	70

Какие курсы следует выбрать студенту?

5. У меня во дворе имеется небольшой огород 10×20 футов. Этой весной я собираюсь посадить овощи трех видов: помидоры, зеленые бобы и кукурузу. Огород разбит на ряды, длина которых равна 20 футам. Кукуруза и помидоры занимают ряды шириной 2 фута, а зеленые бобы — 3 фута. Помидоры мне нравятся больше, а бобы меньше. По 10-балльной шкале предпочтений я бы присвоил помидорам 10 баллов, кукурузе — 7 баллов и зеленым бобам — 3 балла. Независимо от моих предпочтений, жена настаивает, чтобы я посадил не менее одного ряда зеленых бобов и не более двух рядов помидоров. Сколько рядов каждого вида овощей следует мне посадить?

6. "Жилище для Человечества" — прекрасная благотворительная организация, которая строит дома для бедствующих семей силами добровольцев. Такая семья может выбрать себе дом из трех типоразмеров: 1000, 1100 и 1200 квадратных футов. Дом каждого типоразмера требует выполнения определенного объема работ силами добровольцев. Филиал организации в городе Файтвилл получил пять заявок на предстоящие шесть месяцев. Комитет по надзору дает оценку каждой заявке в численном виде, принимая во внимание различные факторы. Более высокая оценка означает более острую потребность в жилье. В

течение предстоящих шести месяцев филиал организации в этом городе может привлечь к работе максимум 23 добровольца. Следующая таблица содержит оценку каждой заявки и необходимое число добровольцев для ее выполнения. Какие заявки следует утвердить комитету?

Заявка	Размер дома (футов^2)	Оценка	Необходимое число добровольцев
1	1200	78	7
2	1000	64	4
3	1100	68	6
4	1000	62	5
5	1200	85	8

7. Шериф округа Вашингтон принимает участие в переизбрании на следующий срок. Денежные средства на предвыборную кампанию составляют примерно 10000 долларов. Хотя комитет по переизбранию хотел бы провести кампанию во всех пяти избирательных участках округа, ограниченность денежных средств предписывает действовать по-другому. Приведенная ниже таблица содержит данные о числе избирателей и денежных средствах, необходимых для проведения успешной кампании по каждому избирательному участку. Каждый участок может либо использовать все предназначенные деньги, либо вовсе их не использовать. Как следует распределить денежные средства?

Участок	Число избирателей	Необходимые средства ($)
1	3100	3500
2	2600	2500
3	3500	4000
4	2800	3000
5	2400	2000

8. Конструируется электронный прибор, состоящий из трех основных компонентов. Все компоненты соединены последовательно, поэтому выход из строя одного из них влечет за собой отказ всего прибора. Надежность (вероятность безаварийной работы) прибора можно повысить путем дублирования каждого компонента. Конструкция прибора допускает использование одного или двух резервных (параллельных) блоков, т.е. каждый компонент прибора может содержать до трех блоков, соединенных параллельно. Следующая таблица содержит данные о надежности r и стоимости компонентов прибора.

Число параллельных блоков	Компонент 1		Компонент 2		Компонент 3	
	r_1	c_1($)	r_2	c_2($)	r_3	c_3($)
1	0,6	1000	0,7	3000	0,5	2000
2	0,8	2000	0,8	5000	0,7	4000
3	0,9	3000	0,9	6000	0,9	5000

Общая сумма, выделенная на конструирование прибора, равна 10000 долларов. Как следует сконструировать прибор? *(Совет.* Наша задача состоит в максимизации надежности $r_1 r_2 r_3$ прибора. Это значит, что целевая функция является мультипликативной, а не аддитивной.)

9. Решите следующую задачу с помощью метода динамического программирования.

Максимизировать $z = y_1 y_2 ... y_n$

при условиях

$$y_1 + y_2 + ... + y_n = c,$$
$$y_i \geq 0, i = 1,2,...,n.$$

(Подсказка это упражнение аналогично предыдущему упражнению, но с той лишь разницей, что переменные y_i являются непрерывными)

10. Решите следующую задачу с использованием метода динамического программирования

Минимизировать $z = y_1^2 + y_2^2 + ... + y_m^2$

при условиях

$$y_1 y_2 ... y_n = c,$$
$$y_i \geq 0,\ \ i = 1,2,...,n.$$

11. Решите следующую задачу посредством метода динамического программирования.

Максимизировать $z = (y_1 + 2)^2 + y_2 y_3 + (y_4 - 5)^2$

при условиях

$$y_1 + y_2 + y_3 + y_4 \leq 5$$
$$y_i \geq 0 \ \textit{и целые},\ \ i = 1,2,3,4.$$

12. Решите следующую задачу с помощью метода динамического программирования.

Минимизировать $z = \max\{f(y_1), f(y_2), ..., f(y_n)\}$

при условиях

$$y_1 + y_2 + ... + y_n = c,$$
$$y_i \geq 0, i = 1,2,...,n.$$

Найдите решение задачи при условии, что $n = 3, c = 10,\ f(y_1) = y_1 + 5,\ f(y_2) = 5y_2 + 3$ и $f(y_3) = y_3 - 2.$

4.4.2. Задача планирования рабочей силы

При выполнении некоторых проектов число рабочих, необходимых для выполнения какого-либо проекта, регулируется путем их найма и увольнения. Поскольку как наем, так и увольнение рабочих связано с дополнительными затратами, необходимо определить, каким образом должна регулироваться численность рабочих в период реализации проекта.

Предположим, что проект будет выполняться в течение *n* недель и минимальная потребность в рабочей силе на протяжении i-й недели составит b_i рабочих. При идеальных условиях хотелось бы на протяжении i-й недели иметь в точности b_i рабочих. Однако в зависимости от стоимостных показателей может быть более выгодным отклонение численности рабочей силы как в одну, так и в другую сторону от минимальных потребностей. Если x_i — количество работающих на протяжении i-й недели, то возможны затраты двух видов: 1) $C_1(x_i - b_i)$ — затраты, связанные с необходимостью содержать избыток x_i - b_i рабочей силы и 2) $C_2(x_i - x_{i-1})$ — затраты, связанные с необходимостью дополнительного найма $x_i - x_{i-1}$ рабочих.

Элементы модели динамического программирования определяются следующим образом.

1. *Этап i* представляется порядковым номером недели $i, i = 1,2,...,n.$

2. Вариантами решения на i-м этапе являются значения x_i — количество работающих на протяжении i-й недели.

3. *Состоянием* на i-м этапе является x_{i-1} — количество работающих на протяжении $(i-1)$й недели (этапа).

Рекуррентное уравнение динамического программирования представляется в виде

$$f_i(x_{i-1}) = \min_{x_i \geq b_i}\{C_1(x_1 - b_i) + C_2(x_i - x_{i-1}) + f_{i+1}(x_i)\}, i = 1,2,...,n,$$

где. Вычисления начинаются с этапа n при $x_n = b_n$ и заканчиваются на этапе 1.

Пример 4-2

Строительный подрядчик оценивает минимальные потребности в рабочей силе на каждую из последующих пяти недель следующим образом: 5, 7, 8, 4 и 6 рабочих соответственно. Содержание избытка рабочей силы обходится подрядчику в 300 долларов за одного рабочего в неделю, а наем рабочей силы на протяжении одной недели обходится в 400 долларов плюс 200 долларов за одного рабочего в неделю.

Выражая C_1 и C_2 в сотнях долларов, имеем следующее.

$$b_1 = 5, b_2 = 7, b_3 = 8, b_4 = 4, b_5 = 6,$$

$$C_1(x_i - b_i) = 3(x_i - b_i), x_i > b_i, i = 1,2,...,5,$$

$$C_2(x_i - x_{i-1}) = 4 + 2(x_i - x_{i-1}), x_i > x_{i-1}, i = 1,2,...,5$$

Этап 5. $(b_5 = 6)$.

	$C_1(x_5 - 6) + C_2(x_5 - x_4)$	Оптимальное решение	
x_4	$x_5 = 6$	$f_5(x_4)$	x_5
4	3(0)+4+2(2)=8	8	6
5	3(0)+4+2(1)=6	6	6
6	3(0)+0=0	0	6

Этап 4. $(b_4 = 4)$

	$C_1(x_4 - 4) + C_2(x_4 - x_3) + f_5(x_4)$			Оптимальное решение	
x_3	$x_4 = 4$	$x_4 = 5$	$x_4 = 6$	$f_5(x_4)$	x_4 *
8	3(0)+0+8=8	3(1)+0+6=9	3(2)+0+0=6	6	6

Этап 3. $(b_3 = 8)$

	$C_1(x_3 - 8) + C_2(x_3 - x_2) + f_4(\tilde{o}_3)$	Оптимальное решение	
x_2	$x_3 = 8$	$f_3(x_2)$	x_3 *
7	3(0)+4+2(1)+6=12	12	8
8	3(0)+0+6=6	6	8

Этап 2. $(b_2 = 7)$

	$C_1(x_2 - 7) + C_2(x_3 - x_2) + f_3(x_2)$		Оптимальное решение	
x_1	$x_2 = 7$	$x_2 = 8$	$f_2(x_1)$	x_2 *
5	3(0)+4+2(2)+12=20	3(1)+4+2(3)+6=19	19	8
6	3(0)+4+2(1)+12=18	3(1)+4+2(2)+6=17	17	8
7	3(0)+0+12=12	3(1)+4+2(1)+6=15	12	7
8	3(0)+0+12=12	3(1)+4+2=9	9	8

Этап 1. $(b_1 = 5)$

x_0	$C_1(x_1-5)+C_2(x_1-x_0)+f_2(x_1)$				Оптимальное решение	
	$x_1=5$	$x_1=6$	$x_1=7$	$x_1=8$	$f_1(x_0)$	x_1*
0	3(0)+4+2(5) +19=33	3(1)+4+2(6) +17=36	3(2)+4+2(7) +12=36	3(2)+4+2(8) +9=35	33	5

Оптимальное решение определяется последовательно следующим образом.

$$x_0=0 \to x_1*=5 \to x_2*=8 \to x_3*=8 \to x_4*=6 \to x_5*=6.$$

Полученному решению соответствует следующий план.

Номер недели (i)	Минимум рабочей силы (b_i)	Количество фактически работающих (x_i)	Решение
1	5	5	Нанять 5 раочих
2	7	8	Нанять 3 рабочих
3	8	8	Ничего не менять
4	4	6	Уволить 2 рабочих
5	6	6	Ничего не менять

Упражнения 4. 4,b

1. Решите задачу из примера 4-2 при следующих минимальных потребностях в рабочей силе.

a) $b_1=6, b_2=5, b_3=3, b_4=6, b_5=8.$

b) $b_1=8, b_2=4, b_3=7, b_4=8, b_5=2.$

2. Пусть в примере 4-2 каждому уволенному рабочему выплачивается выходное пособие в размере 100 долларов. Найдите оптимальное решение задачи.

3. Туристическое агентство организовывает недельные поездки в Египет. В соответствии с договором на ближайшие четыре недели агентство должно обеспечить туристические группы арендными автомобилями в количестве семь, четыре, семь и восемь штук соответственно. Агентство заключает договор с местным дилером по прокату автомобилей. Дилер назначает арендную плату за один автомобиль 220 долларов в неделю плюс 500 долларов за любую арендную сделку. Агентство, однако, может не возвращать арендованные автомобили в конце недели, и в этом случае оно должно будет только арендную плату в 220 долларов. Каково оптимальное решение проблемы, связанной с арендой автомобилей?

4. Компания на следующие четыре года заключила контракт на поставку авиационных двигателей, по 4 двигателя в год. Доступные производственные мощности и стоимость производства меняются от года к году. Компания может изготовить пять двигателей за 1-й год, шесть— за 2-й, три— за 3-й и пять— за 4-й. Стоимость производства одного двигателя на протяжении следующих четырех лет равна соответственно 300 000, 330 000, 350 000 и 420 000 долларов. В течение года компания может произвести больше двигателей, чем необходимо, но в этом случае двигатели должны надлежащим образом храниться до их отгрузки потребителю. Стоимость хранения одного двигателя также меняется от года к году и оценивается в 20 000 долларов для первого года, 30 000 долларов — для второго, 40 000 долларов — для третьего и 50 000 долларов — для четвертого. В начале первого года компания имеет один двигатель, готовый к отгрузке. Разработайте оптимальный план производства двигателей.

5. Фирма выпускает пять типов электронных игр (*E1, E2, ..., E5*) и пять типов механических игрушек (*Ml, M2,,..., M5*). На рынке порядок предпочтения электронных игр таков: $E1 \to E2 \to ... \to E5$. Это означает, что клиент будет покупать игру с более высоким предпочтением, если она имеется в продаже. Известен также порядок предпочтения механических игрушек: $M1 \to M2 \to ... \to M5$. Недельный спрос на пять типов электронных игр равен 100, 180, 90, 250 и 190 единиц соответственно. Аналогичные показатели для механических игрушек равны 300, 190, 240, 280 и 260 единиц соответственно. Производство одной игры *E1,E2,...,E5* обходится в 10, 12, 8, 9 и 6 долларов соответственно. Изготовление же одной игрушки *Ml,Ml,...,M5* обходится фирме в 4, 5, 3, 2 и 3 доллара соответственно. Организация производства каждой электронной игры или игрушки обходится в 500 долларов. Определите оптимальный план производства игрушек.

4.4.3. Задача замены оборудования

Чем дольше механизм эксплуатируется, тем выше затраты на его обслуживание и ниже его производительность. Когда срок эксплуатации механизма достигает определенного уровня, может оказаться более выгодной его замена. Задача замены оборудования, таким образом, сводится к определению оптимального срока эксплуатации механизма.

Предположим, что мы занимаемся заменой механизмов на протяжении *n* лет. В нчале каждого года принимается решение либо об эксплуатации механизма еще один год, либо о замене его новым. Обозначим через $r(t)$ и $c(t)$ и прибыль от эксплуатации $t-$летнего механизма на протяжении года и затраты на его обслуживание за этот же период. Далее пусть *s(t)* — стоимость продажи механизма, который эксплуатировался *t* лет. Стоимость приобретения нового механизма остается неизменной на протяжении всех лет и равна *I*. Элементы модели динамического программирования таковы.

1. *Этап i* представляется порядковым номером года *i,i*= 1,2,...,*n.*

2. Вариантами решения на *i*-м этапе (т.е. для *i*-го года) являются альтернативы: *продолжить эксплуатацию* или *заменить* механизм в *начале i*-го года.

3. *Состоянием* на *i*-м этапе является срок эксплуатации *t* (возраст) механизма к началу *i*-го года.

Пусть $f_i(t)$ — максимальная прибыль, получаемая за годы от *i* до *n* при условии, что в начале *i*-го года имеется механизм *t*-летнего возраста.

Рекуррентное уравнение имеет следующий вид.

$$f_i(t) = \max \begin{cases} r(t) - c(t) + f_{i+t}(t+1) \\ r(0) + s(t) - I - c(0) + f_{i+1}(1), \end{cases}$$

где $f_{i+1}(.) \equiv 0.$

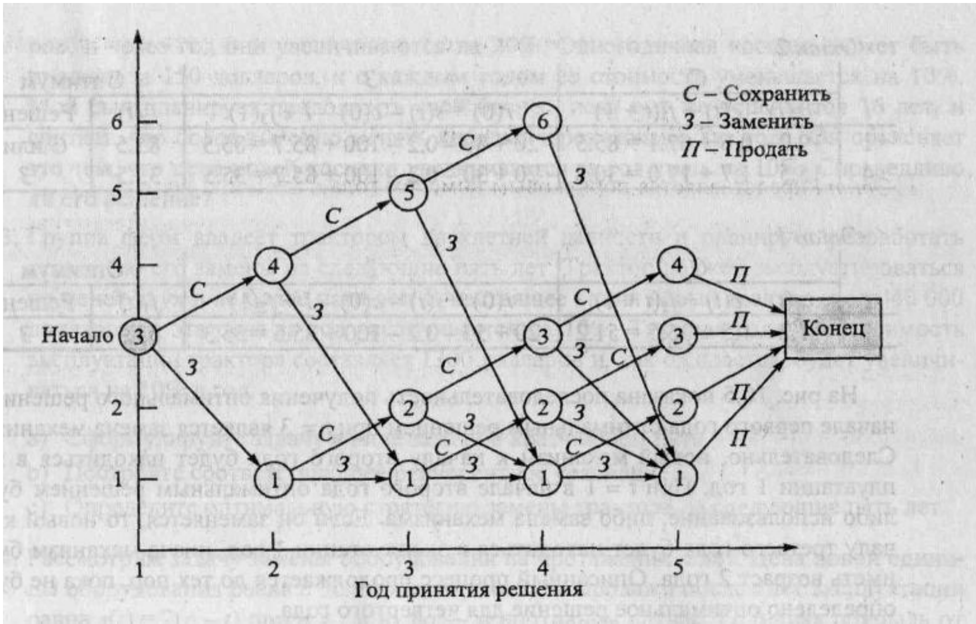
Пример 4-3

Компания планирует определить оптимальную политику замены имеющегося в настоящее время трехлетнего механизма на протяжении следующих 4 лет (*n*=4), т.е. вплоть до *начала* пятого года. Приведенная таблица содержит относящиеся к задаче данные. Компания требует замены механизма, который в эксплуатации 6 лет. Стоимость нового механизма равна 100000 долларов.

Возраст t (года)	Прибыль $r(t)$ ($)	Стоимость обслуживания $c(t)$ ($)	Остаточная стоимость *s(t)* ($)
0	20000	200	–
1	19000	600	80000
2	18500	1200	60000

3	17000	1500	50000
4	15500	1700	30000
5	14000	1800	10000
6	12200	2200	5000

Определение допустимых значений возраста механизма на каждом этапе является нетривиальной задачей. На рис 4 представлена рассматриваемая задача замены оборудования в виде сети. В начале первого года имеется механизм трехлетнего возраста. Мы можем либо заменить его (З), либо эксплуатировать (С) на протяжении следующего года. Если механизм заменили, то в начале второго года его возраст будет равен одному году, в противном случае его возраст будет 4 года. Такой же подход используется в начале каждого года, начиная со второго по четвертый.



Если однолетний механизм заменяется в начале второго или третьего года, то заменивший его механизм к началу следующего года также будет однолетним. К тому же, в начале 4-го года 6-летний механизм обязательно должен быть заменен, если он еще эксплуатируется; в конце 4-го года все механизмы продаются (*П*) в обязательном порядке. На схеме сети также видно, что в начале второго года возможны только механизмы со сроком эксплуатации 1 или 4 года. В начале третьего года механизм может иметь возраст 1, 2 или 5 лет, а в начале четвертого — 1,2,3 или 6 лет.

Решение данной задачи эквивалентно нахождению маршрута максимальной длины (т.е. приносящего максимальную прибыль) от начала первого года к концу четвертого в сети, показанной на рис. 4. При решении этой задачи используем табличную форму записи. (Числовые Данные в таблице кратны тысячам долларов.).

Этап 4.

	С	З	Оптимум	
t	$r(t)+s(t+1)-c(t)$	$r(0)+s(t)+s(1)-c(0)-I$	$f_4(t)$	Решение
1	19.0+60-0.6=78.4	20+80+80-0.2-100=79.8	79.8	*З*
2	18.5+50-1.2=67.3	20+60+80-0.2-100=59.8	67.3	*С*
3	17.2+30-1.5=45.7	20+50+80-0.2-100=49.8	49.8	*З*

6	Необходима замена	20+5+80-0.2-100=4.8	4.8	*З*

Этап 3.

	С	З	Оптимум	
t	$r(t)-c(t)+f_4(t+1)$	$r(0)+s(t)-c(0)-I+f_4(1)$	$f_3(t)$	Решение
1	19.0-0.6+67.3=85.7	20+80-0.2-100+79.8=79.6	85.7	*С*
2	18.5-1.2+49.8=67.1	20+60-0.2-100+79.8=59.6	67.1	*С*
3	14.0+1.8-4.8=17.0	20+10-0.2-100+79.8=9.6	17.0	*С*

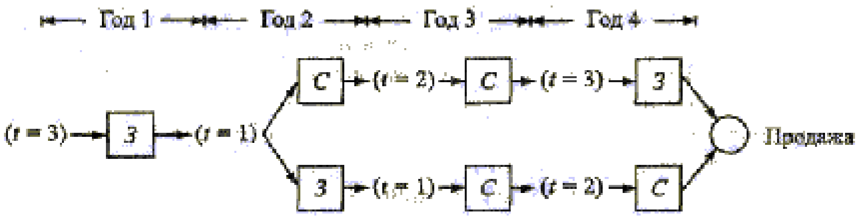
Этап 2.

	С	З	Оптимум	
t	$r(t)-c(t)+f_3(t+1)$	$r(0)+s(t)-c(0)-I+f_3(1)$	$f_2(t)$	Решение
1	19.0-0.6+67.1=85.5	20+80-0.2-100+85.7=85.5	85.5	*С* или *З*
4	19.0-0.6+67.3=85.7	20+80-0.2-100+79.8=79.6	85.7	*З*

Этап 2.

	С	З	Оптимум	
T	$r(t)-c(t)+f_2(t+1)$	$r(0)+s(t)-c(0)-I+f_2(1)$	$f_1(t)$	Решение
1	17.2-1.5+35.5=51.2	20+50-0.2-100+85.5=55.3	55.3	*З*

На рис. 5 показана последовательность получения оптимального решения. В начале первого года оптимальным решением при t = 3 является замена механизма. Следовательно, новый механизм к началу второго года будет находиться в эксплуатации 1 год. При t = 1 в начале второго года оптимальным решением будет либо использование, либо замена механизма. Если он заменяется, то новый к началу третьего года будет находиться в эксплуатации 1 год, иначе механизм будет иметь возраст 2 года. Описанный процесс продолжается до тех пор, пока не будет определено оптимальное решение для четвертого года.



Следовательно, начиная с первого года эксплуатации механизма, альтернативными оптимальными стратегиями относительно замены механизма будут *(З, С, С, З)* и *(З, З, С, С)*. Общая прибыль составит 55 300 долларов.

Упражнения 4.4,с

1. Постройте сеть и найдите оптимальное решение в задаче из примера 4-3 в каждом из следующих случаев.

a) В начале первого года имеется механизм, находящийся в эксплуатации 2 года.

b) В начале первого года имеется механизм, находящийся в эксплуатации 1 год.

c) В начале первого года куплен новый механизм.

2. Мой тринадцатилетний сын занимается собственным бизнесом — косит газоны десяти клиентам. Каждому клиенту он косит траву три раза в год, получая за один скошенный газон 50 долларов. Он купил косилку за 200 долларов. На протяжении первого года затраты на содержание и использование косилки равны 120 долларов, и через год они увеличиваются на 20%. Одногодичная косилка может быть продана за 150 долларов, и с каждым годом ее стоимость уменьшается на 10%. Мой сын планирует продолжить свой бизнес, пока ему не исполнится 16 лет, и считает, что более выгодно менять косилку через

каждые два года. Он объясняет это тем, что цена новой косилки увеличивается за год лишь на 10%. Справедливо ли его решение?

3. Группа ферм владеет трактором двухлетней давности и планирует разработать стратегию его замены на следующие пять лет. Трактор должен эксплуатироваться не менее двух и не более пяти лет. В настоящее время новый трактор стоит 40 000 долларов, и эта цена за год увеличивается на 10%. Текущая годичная стоимость эксплуатации трактора составляет 1300 долларов и, как ожидается, будет увеличиваться на 10% в год.

a) Сформулируйте задачу в виде задачи о кратчайшем пути.

b) Постройте соответствующее рекуррентное уравнение.

c) Определите оптимальную стратегию замены трактора на следующие пять лет.

4. Рассмотрим задачу замены оборудования на протяжении n лет. Цена новой единицы оборудования равна c долларов, а стоимость продажи после t лет эксплуатации равна $s(t)=2(n-t)$ при $n>t$ и нулю — в противном случае. Годичная прибыль от эксплуатации является функцией возраста оборудования t и равна $r(t)=r^2-t^2$ при $n>t$ и нулю — в противном случае.

a) Сформулируйте задачу как модель динамического программирования.

b) Определите оптимальную стратегию замены оборудования двухгодичной давности при c=10 000 долларов, считая, что n=5.

5. Решите задачу из предыдущего упражнения, предполагая, что возраст оборудования составляет 1 год и n = 4, c = 6000 долларов, $r(t) = n/(n + 1)$.

4.4.4. Задача инвестирования

Предположим, что в начале каждого из следующих n лет необходимо сделать инвестиции $P_1, P_2, ..., P_n$ соответственно. Вы имеете возможность вложить капитал в два банка: первый банк выплачивает годовой сложный процент r_1, а второй— r_2. Для поощрения депозитов оба банка выплачивают новым инвесторам премии в виде процента от вложенной суммы. Премиальные меняются от года к году, и для i-го года равны q_{i1} и q_{i2} в первом и втором банках соответственно. Они выплачиваются в конце года, на протяжении которого сделан вклад, и могут быть инвестированы в один из двух банков на следующий год. Это значит, что лишь указанные проценты и новые деньги могут быть инвестированы в один из двух банков. Размещенный в банке вклад должен находиться там до конца рассматриваемого периода. Необходимо разработать стратегию инвестиций на следующие n лет.

Элементы модели динамического программирования следующие.

1. *Этап* i представляется порядковым номером года i, i= 1,2,..., n.

2. *Вариантами решения* на i-м этапе (для i-го года) являются суммы I_i и $\bar{I}_i$, инвестиций в первый и второй банк соответственно.

3. *Состоянием* x_i на i-м этапе является сумма денег на начало i-го года, которые могут быть инвестированы.

Заметим, что по определению $\bar{I}_i = x_i - I_i$. Следовательно,

$$x_i = P_i + q_{i-1}I_{i-1} + q_{i-1}(x_{i-1,2} - I_{i-1}) =$$
$$= P_i + (q_{i-1,1} - q_{i-1,2})I_{i-1} + q_{i-1,2}x_{i-1},$$

где $i = 2,3,...,n, x_1 = P_1$. Сумма денег x_i которые могут быть инвестированы, включает лишь новые деньги и премиальные проценты за инвестиции, сделанные на протяжении $(i-1)$-го года.

Пусть $f_i(x_i)$ — оптимальная сумма инвестиций для интервала от i-го до n-го года при условии, что в начале i-го года имеется денежная сумма x_i. Далее обозначим через s_i - накопленную сумму к концу n-го года при условии, что I_i и (x_i - I_i) — объемы инвестиций на протяжении i -го года в первый и второй банк соответственно. Обозначая $\alpha_i = (1 + r_i),$ i = 1,2, мы можем сформулировать задачу в следующем виде.

Максимизировать $z = s_1 + s_2 + ... + s_n,$

где

$$s_i = I_i \alpha_1^{n+1-i} + (x_i - I_i)\alpha_2^{n+1-i} = (\alpha_1^{n+1-i} - \alpha_2^{n+1-i})I_i + \alpha_2^{n+1-i} x_i, i = 1,2,...n-1,$$

$$s_n = (\alpha_1 + q_{n1} - \alpha_2 - q_{n2} I_n + (\alpha_2 + q_{n2}$$

Так как премиальные за *n-й* год являются частью накопленной денежной суммы от инвестиций, в выражения для s_n добавлены q_{n1} и q_{n2}.

Итак, в данном случае рекуррентное уравнение для обратной прогонки в алгоритме динамического программирования имеет вид

$$f_i(x_i) = \max_{0 \le I_i \le x_i} \{s_i + f_{i+1}(x_{i+1})\}, i = 1,2,...,n-1,$$

где x_{i+1} выражается через x_i в соответствии с приведенной выше формулой, $af_{n+1}(x_{n+1})$ $\equiv 0$.

Пример 4-4

Предположим, вы хотите инвестировать 4000 долларов сейчас и 2000 долларов 3в начале каждого года, от второго до четвертого, считая от текущего года. Первый банк выплачивает годовой сложный процент 8% и премиальные на протяжении следующих четырех лет в размере 1.8%, 1.7%, 2.1% и 2.5% соответственно. Годовой сложный процент, предлагаемый вторым банком, на 0.2% ниже, чем предлагает первый банк, но его премиальные на 0.5% выше. Задача состоит в максимизации накопленного капитала к концу четвертого года.

Используя введенные выше обозначения, имеем следующее.

$$P_1 = \$4000, P_2 = P_3 = P_4 = \$2000,$$

$$\alpha_1 = (1 + 0.08) = 1.08,$$

$$\alpha_2 = (1 + 0.078) = 1.078,$$

$$q_{11} = 0.018, q_{21} = 0.017, q_{31} = 0.021, q_{41} = 0.025,$$

$$q_{12} = 0.023, q_{22} = 0.022, q_{32} = 0.026, q_{42} = 0.030.$$

Этап 4.

$$f_4(x_4) = \max_{0 \le I_4 \le x_4} \{s_4\},$$

где $s_4(\alpha_1 + q_{41} - \alpha_2 - q_{42})I_4 + (\alpha_2 + q_{42})x_4 = -0.003I_4 + 1.108x_4.$

Функция s_4 является линейной по I_4 в области $0<I_4<x_4$ и, следовательно, ее максимум достигается при I_4=0 из-за отрицательного коэффициента при I_4. Следовательно, оптимальное решение для этапа 4 может быть представлено в следующем виде.

Состояние	Оптимальное решение	
	$f_4(x_4)$	I_4*
x_4	1.108x_4	0

Этап 3.

$$f_3(x_3) = \max_{0 \le I_3 \le x_3} \{s_3 + f_4(x_4)\},$$

где

$$s_3 = (1.08^2 - 1.078^2)I_3 + 1.078^2 x_3 = 0.00432I_3 + 1.1621x_3,$$

$$x_4 = 2000 - 0.005I_3 + 0.026x_3.$$

Следовательно,

$$f_3(x_3) = \max_{0\le I_3 \le x_3} \left\{0.00432I_3 + 1.1621x_3 + 1.108(2000 - 0.005I_3 + 0.026x_3)\right\} =$$
$$= \max_{0\le I_3 \le x_3} \left\{2216 - 0.00122I_3 + 1.1909x_3\right\}.$$

Состояние	Оптимальное решение	
	$f_3(x_3)$	I_3*
x_3	2216+1.1909x_3	0

Этап 2.

$$f_2(x_2) = \max_{0\le I_2 \le x_2} \left\{s_2 + f_3(x_3)\right\},$$

где

$$s_2 = (1.08^3 - 1.078^3)I_2 + 1.078^3 x_2 = 0.006985I_2 + 1.25273x_2,$$
$$x_3 = 2000 - 0.005I_2 + 0.022x_2.$$

Следовательно,

$$f_2(x_2) = \max_{0\le I_2 \le x_2} \left\{0.006985I_2 + 1.2527x_2 + 2216 + 1.1909(2000 - 0.005I_2 + 0.022x_2)\right\} =$$
$$= \max_{0\le I_2 \le x_2} \left\{4597.8 - 0.0010305I_2 + 1.27893x_2\right\}.$$

Состояние	Оптимальное решение	
	$f_2(x_2)$	I_2*
x_2	4597.8+1.27996x_2	x_2

Этап 1.

$$f_1(x_1) = \max_{0\le I_1 \le x_1} \left\{s_1 + f_2(x_2)\right\},$$

где

$$s_1 = (1.08^4 - 1.078^4)I_1 + 1.078^4 x_1 = 0.01005I_1 + 1.3504x_1,$$
$$x_2 = 2000 - 0.005I_1 + 0.023x_1.$$

Следовательно,

$$f_1(x_1) = \max_{0\le I_2 \le x_2} \left\{0.1005I_1 + 1.078x_1 + .1909(2000 - 0.005I_1 + 0.023x_1)\right\} =$$
$$= \max_{0\le I_1 \le x_1} \left\{7157.1 - 0.00365I_1 + 1.37984x_1\right\}.$$

Состояние	Оптимальное решение	
	$f_1(x_1)$	I_1*
$x_1 = \$4000$	7157.7+1.3849x_1	\$4000

При вычислениях в обратном направлении получаем следующее.

$$x_2 = 2000 - 0.005\times 4000 + 0.023\times 4000 = \$2072,$$
$$x_3 = 2000 - 0.005\times 2072 + 0.022\times 2072 = \$2035.22,$$
$$x_4 = 2000 - 0.005\times 0 + 0.026\times 2035.22 = \$2052.92,$$

Следовательно, оптимальное решение будет записано следующим образом.

Оптимальное решение	Решение, принимаемое инвестором

I_1*= x_1	Инвестировать $x_1 = \$4000$ в первый банк
I_2*= x_2	Инвестировать $x_2 = \$2072$ в первый банк
I_3*=0	Инвестировать $x_3 = \$2035.22$ во второй банк
I_4*=0	Инвестировать $x_4 = \$2052.92$ во второй банк

Упражнения 4.4, d

1. Решите задачу из примера 10.-4 в предложении, что $r_1 = 0.085, r_2 = 0.008$. Кроме того, пусть $P_1 = \$5000, P_2 = \$4000, P_3 = \$3000, P_4 \$2000$.

2. Некий инвестор с начальным капиталом в 10000 долларов должен решить в конце каждого года, сколько денег истратить и сколько инвестировать. Каждый инвестированный доллар возвращает $\alpha = 1.09$ доллара в конце года приносят удовлетворение, определяемое количественно как эквивалент получения $g(y) = \sqrt{y}$ долларов. Решите задачу с помощью методов динамического программирования для периода в $n = 5$ лет.

3. Фермер имеет k овец. В конце каждого года он принимает решение, сколько овец продать и сколько оставить. Прибыль от продажи одной овцы в i-й год равна p_i. Количество овец в конце i-го года удваивается к концу (i+1)-го года. Фермер планирует в конце n-го года полностью продать овец.

a) Получите общее рекуррентное уравнение для решения задачи.

c) Решите задачу при следующих данных: n=3 года, k=2 овцы, p_1=\$100, p_2 = \$130, p_3=\$120.

d)

4.4.5. Модели управления запасами

Важной областью применения методов динамического программирования являются задачи управления запасами. В главах 11 и 16 рассмотрены некоторые задачи этого класса, при этом в главе 11 рассматриваются детерминированные модели, а в главе 16 — стохастические.

4.5. Проблема размерности

Во всех рассмотренных выше задачах динамического программирования *состояние* системы на любом этапе описывалось единственной переменной. Например, в задаче о загрузке (раздел 10.4.1) вес предмета является единственным ограничением, которое учитывается при его погрузке. Вместе с этим объем судна также может быть ограничительной величиной. В этом случае говорят, что *состояние* системы является двухмерным, так как формируется двумя переменными: весом и объемом.

Увеличение числа переменных состояния системы влечет за собой увеличение объема вычислений на каждом этапе. Особенно это заметно в моделях динамического программирования при вычислениях с использованием таблиц, так как количество строк каждой таблицы должно соответствовать всем возможным комбинациям значений переменных состояния. Эти вычислительные трудности настолько значительны в динамическом программировании, что в литературе на них ссылаются как на проклятие размерности.

Следующий пример приводится для иллюстрации *проклятия размерности.* Он также демонстрирует возможность решения задачи линейного программирования методами динамического программирования.

Пример 4.5-1

Предприятие обрабатывающей промышленности выпускает два вида продукции. Производственный процесс составляет 430 минут в день. Для производства единицы продукции первого вида требуется 2 минуты, а второго — 1 минута. На дневной объем производства продукции первого вида ограничений нет (кроме возможностей производственного процесса), максимальный ежедневный спрос на второй вид продукции равен 230 единиц. Реализация единицы продукции первого вида приносит прибыль в 2 доллара, а второго — 5 долларов. Необходимо найти оптимальное решение задачи максимизации прибыли методами динамического программирования.

Данная задача является следующей задачей линейного программирования.

Максимизировать z = $2x_1 + 5x_2$

при ограничениях

$$2x_1 + x_2 \le 430,$$
$$x_2 \le 230,$$
$$x_1, x_2 \ge 0.$$

Элементы модели динамического программирования таковы.

1. *Этап i* соответствует продукции i, i - 1,2.
2. *Альтернативой* x_i на i-м этапе является объем производства продукции i, i=l,2.
3. *Состояние* представляет количество ресурсов, необходимое для производства продукции вида 1 и 2 (производственное время и ограничение на спрос) и используемое на этапах 1 и 2.
4. *Состояние* (v_2, w_2) представляет количество ресурсов, необходимое для производства продукции вида 1 и 2 (производственное время и ограничение на спрос) и используемое на этапе 2.

Этап 2.

Пусть $f_2(v_2, w_2)$ представляет максимальную прибыль для этапа 2 (прибыль от выпуска продукции вида 2) при заданном состоянии (v_2, w_2). Тогда

$$f_2(v_2, w_2) = \max_{\substack{0 \le x_2 \le v_2 \\ 0 \le x_2 \le w_2}} \{5x_2\}$$

Следовательно, $\max\{5x_2\}$ имеет место при $x_2 = \min\{v_2, w_2\}$. Имеем следующее решение для второго этапа.

Состояние	Оптимальное решение	
	$f_2(x_2)$	x_2
(v_2, w_2)	$5\min\{v_2, w_2\}$	$\min\{v_2, w_2\}$

Этап 1.

$$f_1(v_1, w_1) = \max_{0 \le 2x_1 \le x_1} \{2x_1 + f_2(v_1 - 2x_1, w_1)\} = \max_{0 \le 2x_1 \le v_1} \{2x_1 + 5\min(v_1 - 2x_1, w_1)\}.$$

Оптимизация па первом этапе требует решение минимаксной задачи, что в общем случае является достаточно сложным делом. Для рассматриваемой задачи имеем $v_1 = 430$ и $w_1 = 430$, что дает $0 \le 2x_1 \le 430$. Так как $\min(430 - 2x_1, 230)$ представляет собой нижнюю огибающую двух пересекающихся прямых (проверьте!), отсюда следует, что

$$\min(430 - 2x_1, 230) = \begin{cases} 230, & 0 \le x_1 \le 100, \\ 430 - 2x_1 & 100 \le x_1 \le 215. \end{cases}$$

Графически можно легко проверить, что функция $f_1(430, 230)$ достигает максимального значения при $x_1 = 100$. Итак получаем следующее:

Состояние	Оптимальное решение

	$f_1(v_1, w_1)$	x_1
(430,230)	1350	100

Для определения оптимального значения x_2 заметим, что

$$v_2 = v_1 - 2x_1 = 430 - 200 = 230,$$

$$w_2 = w_1 - 0 = 230.$$

Следовательно

$$x_2 = \min\{v_2, w_2\} = 230.$$

Итак, оптимальное решение имеет вид:

$x_1 = 100$ единиц, $x_2 = 230$ единиц, $z = 1350$.

Упражнения 4.5,а

1. Решите следующие задачи линейного программирования методами динамического программирования.

a) Максимизировать $z = 4x_1 + 14x_2$

при ограничениях

$$2x_1 + 7x_2 \le 8,$$

$$7x_1 + 2x_2 \le 21,$$

$$x_1, x_2 \ge 0.$$

b) Максимизировать $z = 8x_1 + 7x_2$

при ограничениях

$$2x_1 + x_2 \le 8,$$

$$5x_1 + 2x_2 \le 15$$

$$x_1, x_2 \ge 0 \text{ } \textit{и целые}$$

c) Максимизировать $z = 7x_1^2 + 6x_1 + 5x_2^2$

при ограничениях

$$x_1 + 2x_2 \le 10,$$

$$x_1 - 3x_2 \le 9,$$

$$x_1, x_2 \ge 0.$$

2. Пусть в задаче из примера 4.4-1 о загрузке предметов *n* наименований ограничения самолета по весу и объему представлены величинами *W* и *V* соответственно. Величины w_i, v_i и r_i представляют соответственно вес, объем и прибыль, отнесенные к одному предмету наименования. Необходимо записать рекуррентное уравнение обратной прогонки для алгоритма динамического программирования решения сформулированной задачи.

4.6. Заключение

В этой главе рассмотрены некоторые примеры детерминированных моделей динамического программирования. Принцип оптимальности является основой поэтапного решения задачи динамического программирования. Хотя этот принцип не содержит информации о способах решения подзадач на каждом этапе, его применение существенно облегчает решение многих сложных задач, которые нельзя решить другими методами.

Литература

Bertsekas D. *Dynamic Programming: Deterministic and Stochastic Models*, Prentice Hall, Upper Saddle River, N. J., 1987.

Denardo E. *Dynamic Programming Theory and Application*, Prentice Hall, Upper Saddle River, N. J., 1982.

Dreyfus S., Law A. *The Art and Theory of Dynamic Programming*, Academic Press, N. Y., 1977.

Sniedovich M. *Dynsmic Programming*, Marcel Dekker, N. Y., 1991.

Комплексная задача

• 4-1. Компания проверяет состояние оборудования в конце каждого года и на основании этого принимает следующее решение: либо использовать его еще один год, либо заменить. Однако оборудование, которое находилось в эксплуатации три года, подлежит замене в обязательном порядке. Компания планирует разработать стратегию замены имеющегося оборудования на следующие 10 лет. Соответствующая информация содержится в приведенной ниже таблице. Считается, что в начале первого года все оборудование новое.

Год покупки	Стоимость покупки ($)	Стоимость содержания ($) для данного возраста (лет)			Стоимость продажи ($) для данного возраста (лет)		
		0	1	2	1	2	3
1	10000	200	500	600	9000	7000	5000
2	12000	250	600	680	11000	9500	8000
3	13000	280	550	600	12000	11000	10000
4	13500	320	650	700	12000	11500	11000
5	13800	350	590	630	12000	11800	11200
6	14200	390	620	700	12500	12000	11200
7	14800	410	600	620	13500	12900	11900
8	15200	430	670	700	14000	13200	12000
9	15500	450	700	730	15500	14500	13800
10	16000	500	710	720	15800	15000	14500

Глава V
ВЕРОЯТНОСТНОЕ ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ

5.1. Введение

Вероятностное динамическое программирование (ДП) отличается от детерминированного динамического программирования, описанного в главе 10, тем, что состояния и прибыли на каждом этапе являются случайными. Модели вероятностного ДП возникают, в частности, при рассмотрении стохастических моделей управления запасами и в теории марковских процессов принятия решений. Этим двум темам посвящены главы 16 и 19, поэтому в настоящей главе они не рассматриваются. В этой главе описываются некоторые примеры достаточно общего содержания, призванные показать стохастическую природу ДП.

5.2. Азартная игра

Одна из разновидностей игры в русскую рулетку состоит во вращении колеса, на котором по его периметру нанесены n последовательных чисел от 1 до n. Вероятность того, что колесо в результате одного вращения остановится на цифре i, равна p_i,. Игрок платит x долларов за возможность осуществить m вращений колеса. Сам же игрок получает сумму, равную удвоенному числу, которое выпало при *последнем* вращении колеса. Поскольку игра повторяется достаточно много раз (каждая до m вращений колеса), требуется разработать оптимальную стратегию для игрока.

Мы сформулируем задачу в виде модели ДП, используя следующие определения.

1. *Этап* i соответствует i-му вращению колеса, i = 1, 2, ..., m.

2. *Альтернативы* на каждом этапе состоят в следующем — либо покрутить колесо еще раз, либо прекратить игру.

3. *Состояние* системы j на каждом этапе i представляется одним из чисел от 1 до n, которое выпало в результате *последнего* вращения колеса.

Пусть $f_i(j)$ — максимум ожидаемой прибыли при условии, что игра находится на этапе (вращении) i и исходом последнего вращения есть число j Имеем следующее.

$$\begin{pmatrix}\text{ожидаемая прибыль на этапе } i \text{ при условии,}\\ \text{что исходом последнего вращения есть } j\end{pmatrix} = \begin{cases} 2j, & \text{если игра заканчивается,}\\ \sum\limits_{k=1}^{n} p_k f_{i+1}(k), & \text{если игра подолжается.}\end{cases}$$

Рекуррентное уравнение для $f_i(j)$ можно записать следующим образом.

$$f_{m+1}(j) = 2j,$$

$$f_i(j) = \max\begin{cases} \text{конец игры:} & 2j,\\ \text{продолжение игры} & \sum\limits_{k=1}^{n} p_k f_{i+1}(k),\ i = 2,3,...,m,\end{cases}$$

$$f_1(0) = \sum_{k=1}^{n} p_k f_2(k).$$

Обоснование рекуррентного уравнения сводится к следующему. При первом вращении колеса (i = 1) состоянием системы является j = 0, ибо игра только началась. Следовательно, $f_1(0) = p_1 f_2(l) + + ... + p_n f_2(n)$. После выполнения последнего вращения колеса ($i = m$) имеется лишь один выбор — закончить игру независимо от исхода j m – го вращения. Следовательно $f_{m+1}(j) = 2j$.

Рекуррентные вычисления начинаются с f_{m+1}, заканчиваются при $f_1(0)$ и сводятся таким образом к m + 1 вычислительному этапу. Так как $f_1(0)$ представляет собой ожидаемую прибыль от всех m вращений колеса, а игра обходится игроку в x долларов, имеем следующее.

Ожидаемая прибыль = $f_1(0) - x$.

Пример 5.2-1

Предположим, что по периметру колеса русской рулетки расставлены числа от 1 до 5. Вероятности p_i остановки колеса на числе i соответственно равны следующему: $p_1 = 0.3$, $p_2 = 0.25$, $p_3 = 0.2$, $p_4 = 0.15$, $p_5 = 0.1$. Игрок платит 5 долларов за возможность сделать не более четырех вращений колеса. Определим оптимальную стратегию игрока для каждого из четырех вращений и найдем соответствующий ожидаемый выигрыш.

Этап 5.

$$f_5(j) = 2j.$$

Исход 4-го вращения	Оптимальное решение	
j	$f_5(j)$	Решение
1	2	Закончить
2	4	Закончить
3	6	Закончить
4	8	Закончить
5	10	Закончить

Этап 4.

$$\begin{aligned} f_4(j) &= \max\left\{2j,\ p_1 f_5(1) + p_2 f_5(2) + p_3 f_5(3) + p_4 f_5(4) + p_5 f_5(5)\right\} = \\ &= \max\left\{2j,\ 0.3\times 2 + 0.25\times 4 + 0.2\times 6 + 0.15\times 8 + 0.1\times 10\right\} = \\ &= \max\left\{2j, 5\right\}. \end{aligned}$$

Исход 3-го вращения	Ожидаемая прибыль		Оптимальное решение	
j	Закончить	Вращать	$f_4(j)$	Решение
1	2	5	5	Вращать
2	4	5	5	Вращать
3	6	5	6	Закончить
4	8	5	8	Закончить
5	10	5	10	Закончить

Этап 3.

$$\begin{aligned} f_3(j) &= \max\left\{2j,\ p_1 f_4(1) + p_2 f_4(2) + p_3 f_4(3) + p_4 f_4(4) + p_5 f_4(5)\right\} = \\ &= \max\left\{2j,\ 0.3\times 5 + 0.25\times 5 + 0.2\times 6 + 0.15\times 8 + 0.1\times 10\right\} = \\ &= \max\left\{2j, 6.15\right\}. \end{aligned}$$

Исход 2-го вращения	Ожидаемая прибыль		Оптимальное решение	
j	Закончить	Вращать	$f_3(j)$	Решение
1	2	6.15	6.15	Вращать
2	4	6.15	6.15	Вращать
3	6	6.15	6.15	Вращать
4	8	6.15	8.00	Закончить
5	10	6.15	10.00	Закончить

Этап 2.

$$f_2(j)=\max\left\{2j,\ p_1f_3(1)+p_2f_3(2)+p_3f_3(3)+p_4f_3(4)+p_5f_3(5)\right\}=$$
$$=\max\left\{2j,\ 0.3\times6.15+0.25\times6.15+0.2\times6.15+0.15\times8+0.1\times10\right\}=$$
$$=\max\left\{2j,6.8125\right\}.$$

Исход 1-го вращения	Ожидаемая прибыль		Оптимальное решение	
j	Закончить	Вращать	$f_3(j)$	Решение
1	2	6.81	6.81	Вращать
2	4	6.81	6.81	Вращать
3	6	6.81	6.81	Вращать
4	8	6.81	8.00	Закончить
5	10	6.81	10.00	Закончить

Этап.1.

$$f_1(0)=p_1f_2(1)+p_2f_2(2)+p_3f_2(3)+p_4f_2(4)+p_5f_2(5)=$$
$$=0.3\times6.8125+0.25\times6.8125+0.2\times6.8125+0.15\times8+0.1\times10=7.31.$$

В начале игры единственным выбором является вращение колеса.

Из предыдущих таблиц следует, что оптимальным решением будет следующая последовательность действий.

Номер вращения	Оптимальная стратегия
1	Начало игры; вращать
2	Продолжить игру, если исходом первого вращения есть 1, 2 или 3; иначе закончить игру
3	Продолжить игру, если исходом первого вращения есть 1, 2 или 3; иначе закончить игру
4	Продолжить игру, если исходом первого вращения есть 1 или 2; иначе закончить игру

Ожидаемая прибыль от игры составляет 7.31 – 5 = 2.31 доллара.

Упражнение 5.2, а

1. Пусть в задаче из примера 5.2-1 по периметру колеса расставлены числа от 1 до 8 и вероятности остановки колеса на каждом из этих чисел одинаковы. Предполагая, что каждая игра состоит не более чем из пяти вращений колеса, определите оптимальную стратегию игры.

2. Я хочу продать свой подержанный автомобиль тому, кто предложит наивысшую цену. Изучая автомобильный рынок, я пришел к выводу, что с равными вероятностями мне за автомобиль могут предложить очень низкую цена (около 1050 долларов), просто низкую цену (около 1900 долларов), среднюю цену (около 2500 долларов) либо высокую цену (примерно 3000 долларов). Я решил помещать объявление о продаже автомобиля на протяжении не более трех дней подряд. В конце каждого дня мне следует решить, принять ли наилучшее предложение, поступившее в течение этого дня. Какой должна быть моя оптимальная стратегия относительно принятия предложенной цены за автомобиль?

5.3. Задача инвестирования

Некто планирует инвестировать C тысяч долларов через фондовую биржу в течение последующих n лет. Инвестиционный план состоит в покупке акций в начале года и продаже их в конце этого же года. Накопленные деньги затем могут быть снова инвестированы (все или их часть) в начале следующего года. Степень риска инвестиции представлена тем, что прибыль имеет вероятностный характер. Изучение рынка свидетельствует о том, что прибыль от инвестиции зависит от m условий рынка (благоприятных или неблагоприятных). При этом условие i приводит к прибыли r_i с вероятностью p_i, $i = 1, 2, ..., m$. Как следует инвестировать C тысяч долларов для наибольшего накопления к концу n лет?

Обозначим

x_i – сумма денежных средств, доступных для инвестирования в начале i-го года ($x_1=C$),

y_i – сумма реальной инвестиции в начале i-го года ($y_i \le x_i$).

Элементы модели ДП можно описать следующим образом.

1. *Этап* i представляет i-й год инвестирования.
2. *Альтернативами* на этапе i являются величины y_i.
3. *Состояние* системы на этапе i описывается величиной x_i.

Пусть $f_i(x_i)$ — максимальная ожидаемая сумма поступления денежных средств за года от i до n при условии, что в начале i-го года имеется сумма x_i. Для k-го условия рынка имеем следующее.

$$x_{i+1} = (1+r_k)y_i + (x_i - y_i) = r_k y_i + x_i,\ k = 1, 2, ..., m.$$

Так как вероятность k-го условия рынка равна p_k, рекуррентное уравнение динамического программирования имеет следующий вид.

$$f_i(x_i) = \max_{0 \le y_i \le x_i} \left\{ \sum_{k=1}^{m} p_k f_{i+1}(x_i + r_k y_i) \right\},\ i = 1, 2, ..., n,$$

где $f_{n+1}(x_{n+1}) = x_{n+1}$, так как после n-го года инвестиций нет. Отсюда следует, что

$$f_n(x_n) = \max_{0 \le y_n \le x_n} \left\{ \sum_{k=1}^{m} p_k (x_n + r_k y_n) \right\} = x_n \sum_{k=1}^{m} p_k (1+r_k) = x_n (1 + p_1 r_1 + p_2 r_2 + ... + p_m r_m),$$

поскольку функция в фигурных скобках является линейной по y_n и, следовательно, достигает своего максимума при $y_n = x_n$.

Пример 5.3-1

Пусть в предыдущей модели инвестирования объем инвестиции составляет C =10 000 долларов на 4-летний период. Существует 40%-ная вероятность того, что вы удвоите деньги, 20%-ная — останетесь при своих деньгах и 40% — потеряете весь объем инвестиции. Необходимо разработать оптимальную стратегию инвестирования.

Используя принятые в модели обозначения, имеем следующее.

$C = \$10\ 000,\ n = 4,\ m = 3,$
$p_1 = 0.4,\ p_2 = 0.2,\ p_3 = 0.4,$
$r_1 = 2,\ r_2 = 0,\ r_3 = -1.$

Этап 4.

$$f_4(x_4) = x_4(1 + 0.4 \times 2 + 0.2 \times 0 + 0.4 \times (-1)) = 1.4x_4 .$$

Отсюда порлучаем

Состояние	Оптимальное решение	
	$f_4(x_4)$	y_4^*
x_4	$1.4x_4$	x_4

Этап 3.

$$\begin{aligned} f_3(x_3) &= \max_{0 \le y_3 \le x_3} \{p_1 f_4(x_3 + r_1 y_3) + p_2 f_4(x_3 + r_2 y_3) + p_3 f_4(x_3 + r_3 y_3)\} = \\ &= \max_{0 \le y_3 \le x_3} \{0.4 \times 1.4(x_3 + 2y_3) + 0.2 \times 1.4(x_3 + 0y_3) + 0.4 \times 1.4[x_3 + (-1)y_3]\} = \\ &= \max_{0 \le y_3 \le x_3} \{1.4x_3 + 0.56y_3\} = 1.96x_3. \end{aligned}$$

Поэтому имеем

Состояние	Оптимальное решение	
	$f_3(x_3)$	y_3^*
x_3	$1.96x_3$	x_3

Этап 2.

$$\begin{aligned} f_2(x_2) &= \max_{0 \le y_2 \le x_2} \{p_1 f_3(x_2 + r_1 y_2) + p_2 f_3(x_2 + r_2 y_2) + p_3 f_3(x_2 + r_3 y_2)\} = \\ &= \max_{0 \le y_2 \le x_2} \{0.4 \times 1.96(x_2 + 2y_2) + 0.2 \times 1.96(x_2 + 0y_2) + 0.4 \times 1.96[x_2 + (-1)y_2]\} = \\ &= \max_{0 \le y_2 \le x_2} \{1.96x_2 + 0.784y_2\} = 2.744x_2. \end{aligned}$$

Отсюда следует

Состояние	Оптимальное решение	
	$f_2(x_2)$	y_2^*
x_2	$2.744x_2$	x_2

Этап 1.

$$f_1(x_1) = \max_{0 \le y_1 \le x_1} \left\{ p_1 f_2(x_1 + r_1 y_1) + p_2 f_2(x_1 + r_2 y_1) + p_3 f_2(x_1 + r_3 y_1) \right\} =$$

$$= \max_{0 \le y_1 \le x_1} \left\{ 0.4 \times 2.744(x_1 + 2y_1) + 0.2 \times 2.744(x_1 + 0y_1) + 0.4 \times 2.744\left[x_1 + (-1)y_1\right] \right\} =$$

$$= \max_{0 \le y_1 \le x_1} \left\{ 2.744x_1 + 1.0796y_1 \right\} = 3.8416x_1.$$

Имеем

Состояние	Оптимальное решение	
	$f_1(x_1)$	y_1^*
x_1	$2.744x_1$	x_1

Оптимальную инвестиционную политику можно сформулировать следующим образом. Так как $y_i^* = x_i$ для i = 1, 2, 3, 4, то оптимальным решением является инвестирование всех наличных денежных средств в начале каждого года. Накопленные денежные средства к концу четырех лет составят $3.8416x_1$ = 3.8416 3 10000 = 38416 долларов.

Упражнения 5.3,а

1. Определите оптимальную инвестиционную политику в примере 5.3-1 в предположении, что вероятности p_k и прибыли r_k для следующих 4 лет принимают такие значения.

Год						
1	2	1	0.5	0.1	0.4	0.5
2	1	0	– 1	0.4	0.4	0.2
3	4	– 1	– 1	0.2	0.4	0.4
4	0.8	0.4	0.2	0.6	0.2	0.2

2. Камера объемом 10 кубических метров предназначена для хранения изделий трех наименований. Одно изделие наименований 1, 2, 3 занимает соответственно 2, 1 и 3 кубических метра. Вероятности спроса на эти изделия приведены в следующей таблице.

Количество единиц	Вероятность спроса		
	Наименование 1	Наименование 2	Наименование 3
1	0.5	0.3	0.3
2	0.5	0.4	0.2
3	0.0	0.2	0.5
4	0.0	0.1	0.0

Стоимость хранения единицы изделия наименований 1, 2, 3 равна 8, 10 и 15 долларов соответственно. Сколько единиц изделий каждого наименования следует хранить в камере?

3. Фирма с высокотехнологичным производством начала выпуск самых современных суперкомпьютеров в расчете на трехлетний период. Годовой спрос D на новый суперкомпьютер описывается распределением

$$p(D=1) = 0.5,\ p(D=2) = 0.3,\ p(D=3) = 0.2.$$

Производственная мощность завода составляет три суперкомпьютера в год стоимостью 5 миллионов долларов каждый. Количество произведенных за год суперкомпьютеров может не совпадать в точности с объемом спроса. На нереализованный к концу года суперкомпьютер требуются затраты в 1 миллион долларов, связанные с его хранением и содержанием в исправности. Фирма терпит убытки в 2 миллиона долларов, если поставка суперкомпьютера откладывается на один год. Фирма не будет принимать новых заказов позже четвертого года, но будет продолжать выпуск суперкомпьютеров на протяжении пятого года, чтобы выполнить все заказы, оказавшиеся невыполненными к концу четвертого года. Определите оптимальные годичные объемы производства суперкомпьютеров.

4. Компания владеет тремя спортивными центрами в деловой части города. На Пасху популярны велосипедные прогулки на открытом воздухе. В компании имеется восемь велосипедов, которые она может распределить между тремя центрами для их проката в целях максимизации доходов. Спрос на велосипеды и часовая стоимость их аренды зависят от месторасположения центра и характеризуются следующими данными.

Количество велосипедов	Вероятность спроса		
	Центр 1	Центр 2	Центр 3
0	0.10	0.02	0
1	0.20	0.03	0.15
2	0.30	0.10	0.25
3	0.20	0.25	0.30
4	0.10	0.30	0.15
5	0.10	0.15	0.10
6	0	0.05	0.025
7	0	0.05	0.025
8	0	0.05	0
Арендная плата ($/час)	6	7	5

Как компании распределить восемь велосипедов между тремя спортивными центрами?

5.4. Максимизация вероятности достижения цели

В разделе 5.3 рассматривалась задача, связанная с максимизацией ожидаемой прибыли. Иным полезным критерием для рассмотренной задачи является максимизация вероятности достижения определенного уровня дохода. Продемонстрируем этот подход на примере модели инвестирования, которая описана в разделе 5.3.

Используя обозначения из раздела 5.3, оставим без изменения определение *этапа i*, *альтернативы* y_i и *состояния* x_i. Эти модели отличаются только определением критерия; здесь нашей целью является максимизация вероятности достижения некоторой накопленной денежной суммы S по истечении n лет. С этой точки зрения определим функцию $f_i(x_i)$ – вероятность накопления суммы S, если в начале i-го года имеются денежные средства в сумме x_i и для последующих лет i, $i + 1,..., n$ используется оптимальное инвестирование.

Рекуррентное уравнение динамического программирования имеет вид

$$f_n(x_n) \max_{0\le y_n\le x_n}\left\{\sum_{k=1}^{m} p_k P\left\{x_n + r_k y_n \ge S\right\}\right\},$$

$$f_i(x_i) \max_{0\le y_i\le x_i}\left\{\sum_{k=1}^{m} p_k\, f_{i+1}(x_i + r_k y_i)\right\},\ i = 1, 2, ..., n-1.$$

Рекуррентная формула основана на формуле условной вероятности

$$P\{A\} = \sum_{j=1}^{m} P\{A \big| B_j\} P\{B_j\}.$$

В нашем случае f_{i+1} ($x_i + r_k\, y_i$) играет роль вероятности $P\{A \mid B_j\}$.

Пример 5.4-1

Некий индивидуум планирует инвестировать 2 000 долларов. Имеющиеся варианты позволяют удвоить эту сумму с вероятностью 0.3 или потерять ее с вероятностью 0.7. Акции продаются в конце года, а в начале следующего года все деньги или их часть снова инвестируются. Этот процесс повторяется на протяжении трех лет. Целью является максимизация вероятности достижения суммы в 4 000 долларов в конце третьего года.

В соответствии с обозначениями данной модели, имеем $r_1 = 1$ с вероятностью 0.3 и $r_2 = -1$ с вероятностью 0.7.

Этап 3.

На этом этапе состояние x_3 может изменяться от 0 до 8 000 долларов. Минимальное значение возможно, когда вся инвестиция потеряна, а максимальное – когда инвестиция удваивается в конце каждого из двух первых лет. Следовательно, рекуррентное уравнение для этапа 3 записывается в следующем виде:

$$f_3(x_3) = \max_{y_3=0,1,\ldots,x_3} \{0.3P\{x_3 + y_3 \geq 4\} + 0.7P\{x_3 - y_3 \geq 4\}\},$$

где, x_3 = 0, 1,..., 8.

Приведенная ниже таблица содержит детали вычислений для данного этапа. Все заштрихованные ячейки таблицы являются неподходящими, так как не удовлетворяют условию $y_3 \leq x_3$. Кроме того, при выполнении вычислений можно заметить, что

$$\begin{aligned} &P\{x_3 + y_3 \geq 4\} = 0, \text{ если } x_3 + y_3 < 4, \\ &P\{x_3 - y_3 \geq 4\} = 0, \text{ если } x_3 - y_3 < 4. \end{aligned}$$

В противном случае эти вероятности равны 1.

Хотя приведенная таблица и свидетельствует о том, что существуют альтернативные оптимумы для x_3 = 1, 3, 4, *5,* 6, 7 и 8, оптимальный (последний) столбец содержит лишь наименьшие оптимальные значения y_3. Это объясняется тем, что инвестор не собирается инвестировать больше того, что необходимо для достижения поставленной цели.

Этап 2.

$$f_2(x_2) = \max_{y_2=1,2,\ldots x_2} \{0.3 f_3(x_2 + y_2) + 0.7 f_3(x_2 - y_2)\}.$$

x_2	$0.3f_3(x_2+y_2)+0.7f_3(x_2-y_2)$					Оптимум	
	$y_2=0$	$y_2=1$	$y_2=2$	$y_2=3$	$y_2=4$	f_2	y_2
0	$0.3\times 0+0.7\times 0=0$					0	0
1	$0.3\times 0+0.7\times 0=0$	$0.3\times 0.3+0.7\times 0=0.09$				0.09	1
2	$0.3\times 0.3+0.7\times 0.3=0.3$	$0.3\times 0.3+0.7\times 0=0.09$	$0.3\times 1+0.7\times 0=0.3$			0.30	0
3	$0.3\times 0.3+0.7\times 0.3=0.3$	$0.3\times 1+0.7\times 0.3=0.51$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$		0.51	1
4	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0.3=0.51$	$0.3\times 1+0.7\times 0.3=0.51$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	1	0

x_3	$0.3P\{x_3+y_3\ge 4\}+0.7P\{x_3-y_3\ge 4\}$									Оптимум	
	$y_3=0$	$y_3=1$	$y_3=2$	$y_3=3$	$y_3=4$	$y_3=5$	$y_3=6$	$y_3=7$	$y_3=8$	f_3	y_3
0	$0.3\times 0+0.7\times 0=0$									0	0
1	$0.3\times 0+0.7\times 0=0$	$0.3\times 0+0.7\times 0=0$								0	0
2	$0.3\times 0+0.7\times 0=0$	$0.3\times 0+0.7\times 0=0$	$0.3\times 1+0.7\times 0=0.3$							0.3	2
3	$0.3\times 0+0.7\times 0=0$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$						0.3	1
4	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$					1	0
5	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$				1	0
6	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$			1	0
7	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$		1	0
8	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 1=1$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	$0.3\times 1+0.7\times 0=0.3$	1	0

Этап 1.

$$f_1(x_1) = \max_{y_1=1,2,\dots x_1} \{0.3 f_2(x_1+y_1) + 0.7 f_2(x_1-y_1)\}.$$

x_1	$0.3 f_2(x_1+y_1)+0.7 f_2(x_1-y_1)$			Оптимум	
	$y_1 = 0$	$y_1 = 1$	$y_1 = 2$	f_1	y_1
2	0.330.3+0.730.3 = 0.3	0.330.51+0.730.09 = 0.216	0.331+0.730 = 0.3	0.3	0

Оптимальная стратегия определяется следующим образом. При заданной начальной сумме x_1 = 2000 долларов вычисления для первого этапа дают y_1 = 0. Это означает, что в первый год не следует делать инвестиций. Данное решение оставляет инвестора с 2000 долларов к началу второго года. Из таблицы, соответствующей второму этапу, при x_2 = 2 получаем y_2 = 0; это снова означает, что на протяжении второго года также не следует делать инвестиций. Далее использование значения x_3 = 2 на третьем этапе приводит к y_3 = 2, а это означает, что на третий год следует инвестировать всю имеющуюся в распоряжении сумму. Соответствующая максимальная вероятность достижения цели S = 4 равна $f_1(2)$ = 0.3.

Упражнения 5.4,а

1. В примере 5.4–1 этап 1 решения задачи показывает, что существует два альтернативных оптимума: y_1 = 0 и y_1 = 2. Покажите, что применение стратегии y_1 = 2 (т.е. инвестировать все деньги в начале первого года) не изменяет результата инвестиционной политики на протяжении трех лет, а именно, соответствующая максимальная вероятность достижения цели сохраняется равной 0.3.

2. Решите задачу из примера 5.4-1, если целью инвестора является максимизация вероятности достижения, по меньшей мере, суммы в 6 000 долларов к концу третьего года. Инвестор имеет в своем распоряжении 1000 долларов, и вероятность удвоения суммы на протяжении каждого года равна 0.6.

3. Вы и ваш друг хотите сыграть в казино в следующую игру. Вы делаете определенную ставку, и каждый из вас независимо подбрасывает симметричную монету. За каждый доллар суммы ставки казино заплатит три доллара, (что дает чистую прибыль в 2 доллара), если в результате подбрасывания выпадут две решки. Иначе вы теряете сумму ставки. Если вы с другом имеете в сумме один доллар, определите стратегию игры, считая, что целью является максимизация вероятности окончания трех игр с суммой в 4 доллара.

Литература

Bertsekas D. *Dynamic Programming: Deterministic and Stochastic Models*, Prentice Hall, Upper Saddle River, N. J., 1987.

Cooper L. and Cooper M. *Introduction to Dynamic Programming*, Pergamon Press, N. Y., 1981.

Smith D. *Dynamic Programming: A Practical Introduction*, Ellis Horwood, London, 1991.

Комплексные задачи

5-1. Компания использует грузовые автомобили для доставки заказов покупателям и планирует заменить свои автомобили на протяжении последующих пяти лет. Годовые затраты, связанные с использованием нового грузовика, являются нормально распределенной случайной величиной с математическим ожиданием 300 долларов и среднеквадратическим отклонением 50 долларов. Математическое ожидание и среднеквадратическое отклонение годовых эксплуатационных затрат через год возрастают на 10%. Стоимость нового грузового автомобиля в настоящее время равна 20 000 долларов и через год возрастет, как ожидается, на 12%. Грузовые автомобили используются чрезвычайно интенсивно, поэтому существует вероятность того, что каждый из них может сломаться в любое время, после чего ремонту не подлежит. Имеется возможность сдать старый автомобиль при покупке нового. При этом стоимость старого автомобиля зависит от того, находится ли он в рабочем состоянии. В начале шестого года автомобиль подлежит продаже по цене, которая также зависит от его состояния (аварийное или рабочее). Приведенная ниже таблица содержит данные, описывающие ситуацию в зависимости от возраста автомобиля.

Возраст автомобиля (года)	0	1	2	3	4	5	6
Вероятность поломки	0.01	0.05	0.10	0.16	0.25	0.40	0.60

Если автомобиль использовался 1 год и находится в рабочем состоянии, то его стоимость равна 70% от начальной и за год уменьшается на 15%. Если же он находится в аварийном состоянии, то соответствующие показатели уменьшаются в два раза. Стоимость автомобиля в виде испорченного имущества в начале шестого года составляет 200 долларов, если он находится в рабочем состоянии, и 50 долларов, если он аварийный. Разработайте оптимальную политику замены автомобилей.