

Динамическое программирование (планирование) применяется для нахождения оптимальных решений в многошаговых (многоэтапных) задачах.

Принцип оптимальности динамического программирования сформулирован Р.Беллманом:

«Оптимальное поведение обладает тем свойством, что каковы бы ни были первоначальное состояние и решение в начальный момент, последующие решения должны составлять оптимальное поведение относительно состояния, полученного в результате первоначального решения».

Примерами задач, решаемых методом динамического программирования, могут служить:

- планирование производственной программы по периодам года при минимальных затратах на производство и содержание запасов;
- распределение средств и ресурсов на расширение производства при максимизации прироста выпуска продукции;
- оптимальное планирование замены устаревшего оборудования более совершенным при максимизации прибыли;
- календарное планирование ремонта либо замены устаревшего оборудования при минимуме эксплуатационных затрат;
- выбор оптимального маршрута при минимуме транспортных расходов

Любую многоэтапную задачу можно решать двумя способами: искать оптимальное решение сразу на всем протяжении процесса или строить его шаг за шагом.

После оптимизации i -го шага, исходя из результатов предыдущего, оптимизируют следующий $(i + 1)$ -й шаг.

Второй способ проще.

Его суть в постепенной, поэтапной оптимизации, что особенно ценно в задачах, где ситуация меняется во времени, и поэтому необходимо планировать с учетом временного фактора.

Принцип оптимальности Беллмана записывается в виде рекуррентного соотношения

$$f_{n-i}(x_i) = \max_{x_i, x_{i+1}}(\min) \left\{ R_{i+1}(x_i, x_{i+1}) + f_{n-(i+1)}(x_{i+1}) \right\},$$

(i=1, n-1)

где x_i, x_{i+1} — решение (управление), выбранное на i -м шаге -дуга из

$$x_i \text{ в } x_{i+1}$$

$x_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(m)})$ — состояние системы на i -м шаге;

R , — эффект, достигаемый на i -м шаге;

f_{n-i} — оптимальное значение эффекта, достигаемого за $n-i$ шагов;

n число шагов (этапов).

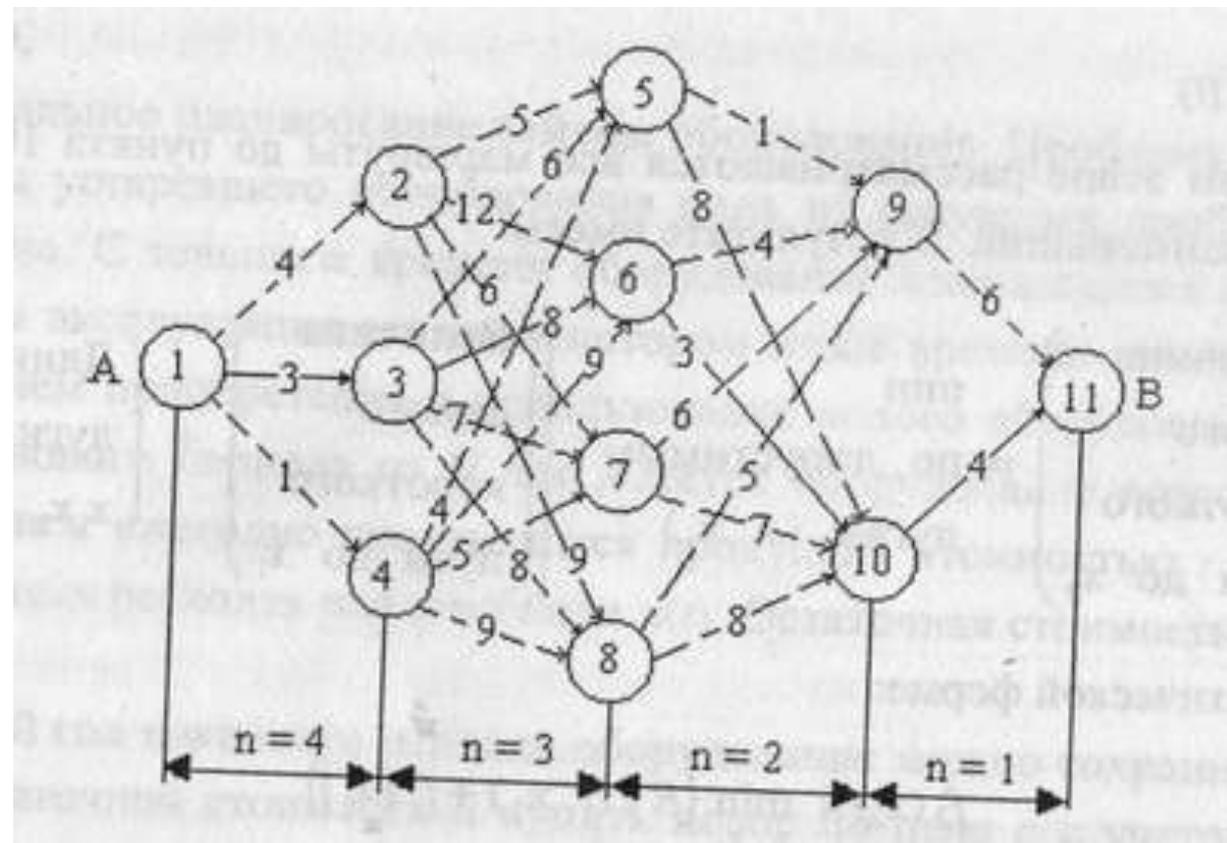
Пример.

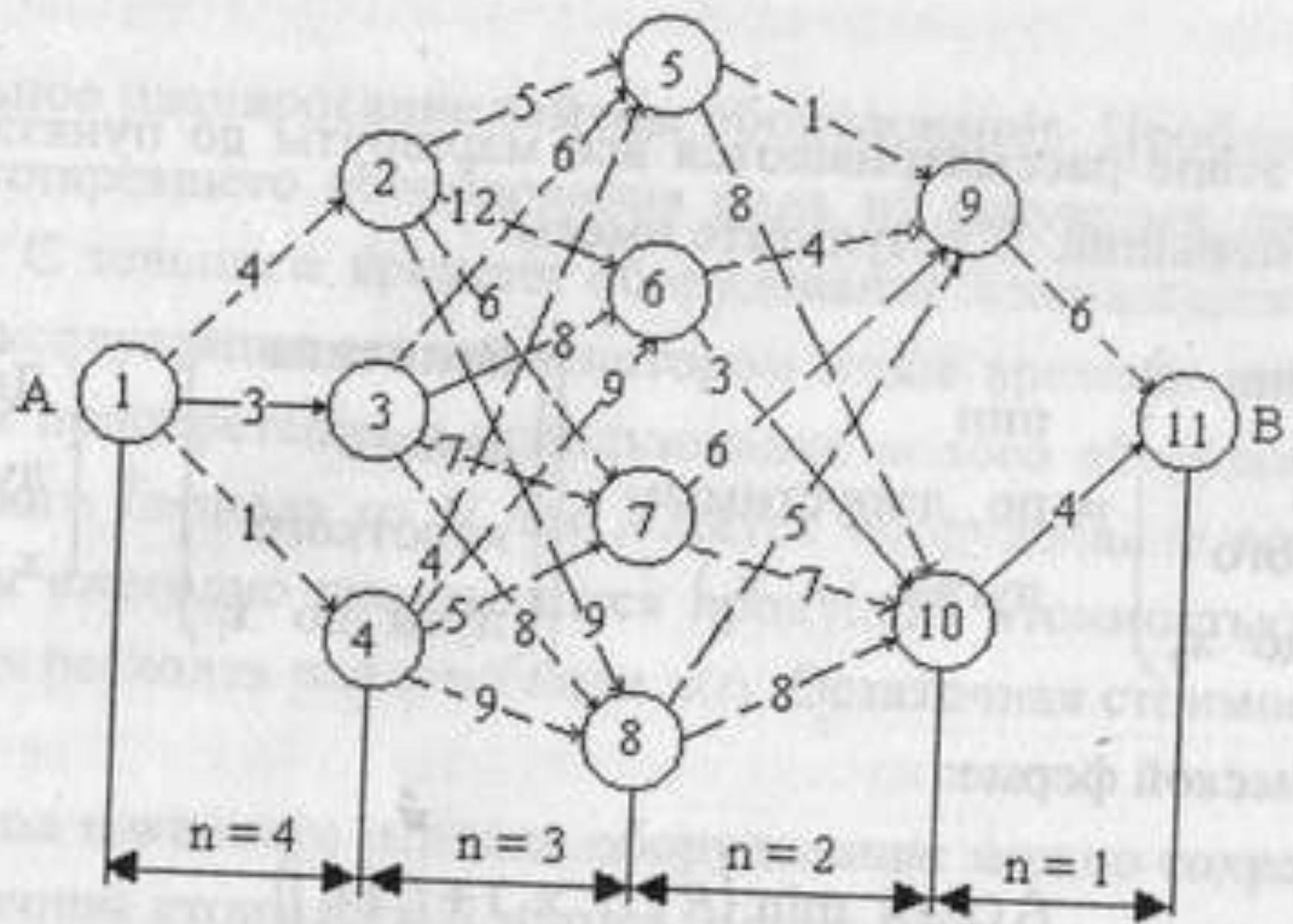
Требуется перевезти груз из пункта А в пункт В. Сеть дорог, связывающих эти пути, показана на рис.

Стоимость перевозок груза из пункта $x_{n-1} (x_{n-1} = \overline{1,10})$

в пункт $x_n (x_n = \overline{2,11})$

поставлена над соответствующими дугами сети. Необходимо найти маршрут, связывающий пункты А В, для которого суммарные транспортные расходы были бы наименьшими.





Решение.

Цифры в вершинах сети означают промежуточные пункты (2 - 10). а дугам соответствуют транспортные магистрали с указанием стоимости (условные единицы) перевозки.

Разобьем все множество пунктов на пять подмножеств (см. рис.): (1), (2,3,4), (5,6,7,8), (9,10), (11). Таким образом, любой маршрут из пункта А в пункт В содержит четыре этапа (четыре дуги).

На первом этапе предстоит решить, через какой пункт, принадлежащий второму подмножеству, предстоит везти груз, затем, на втором этапе — через какой пункт третьего подмножества следует лучший маршрут из некоторого пункта второго множества.

Пронумеруем этапы от конечной вершины сети к начальной (см. рис. 2).

Рассмотрим последний шаг (этап) при $n = 1$ и вычислим для него значение функции :

$$\begin{aligned} f_1(9) &= x_{9,10} + f_0(11) = 6 + 0 = 6; \\ f_1(10) &= x_{10,11} + f_0(11) = 4 + 0 = 4. \end{aligned}$$

Оптимальный маршрут на последнем этапе

$$f_1(x_1) = \min_{x_{n-1}, x_n} \{6; 4\} = 4,$$

т.е. из пункта 10.

На втором этапе рассматриваются все маршруты до пункта 10 и выбирается наименьший. В результате имеем:

$$\left(\begin{array}{l} \text{Величина} \\ \text{самого} \\ \text{короткого} \\ \text{пути до } x_2 \end{array} \right) = \min_{\text{по допустимым} \\ \text{дугам } (x_1, x_2)} \left\{ \left(\begin{array}{l} \text{Величина} \\ \text{самого} \\ \text{короткого} \\ \text{пути до } x_1 \end{array} \right) + \left(\begin{array}{l} \text{Длина} \\ \text{дуги} \\ x_1 x_2 \end{array} \right) \right\},$$

математической форме:

$$f_2(x_2) = \min_{x_1, x_2} \{R(x_1, x_2) + f_1(x_1)\},$$

или в цифрах:

$$f_2(5) = 8 + 4 = 12 \quad f_2(6) = 3 + 4 = 7 \quad f_2(7) = 7 + 4 = 11, \quad f_2(8) = 4 + 4 = 8.$$

Отсюда

$$f_2(x_2) = \min_{x_1, x_2} \{12, 7, 11, 8\} = 7,$$

т.е. на втором этапе лучшим является маршрут из пункта 6 в пункт 10.

На третьем этапе рассматриваются маршруты в пункт 6 и из них выбирается наилучший:

$$f_3(x_3) = \min_{x_2, x_3} \{R_3(x_2, x_3) + f_2(x_2)\}, \quad f_3(3) = 8 + 7 = 15, \quad f_3(4) = 9 + 7 = 16.$$

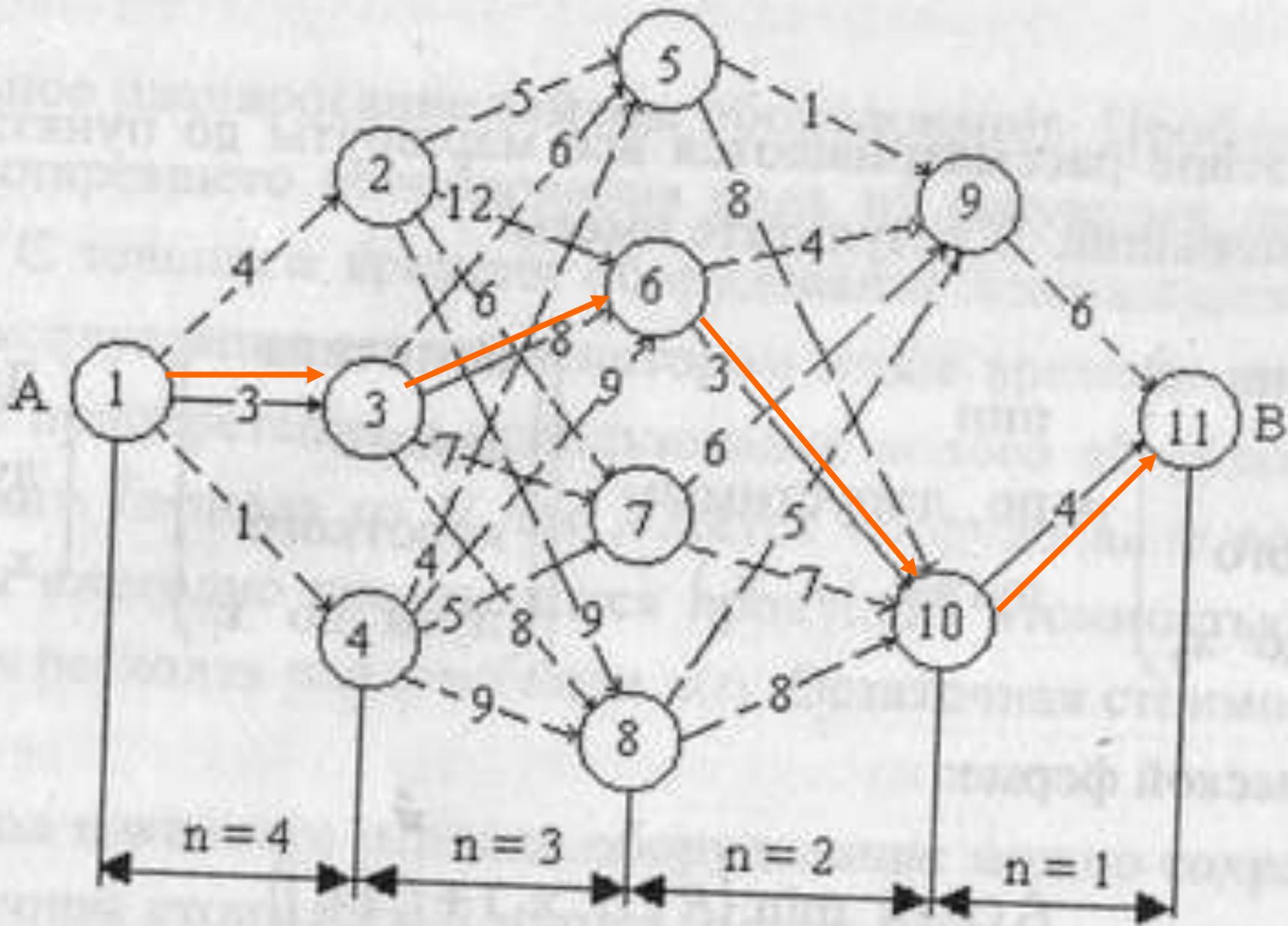
Отсюда

$$f_3(x_3) = \min_{x_2, x_3} \{15; 16\} = 15,$$

т.е. на третьем этапе выбираем маршрут из пункта 3 в пункт 6. И, наконец на четвертом этапе остается один-единственный возможный маршрут

$$f_4(x_4) = 3.$$

Оптимальный маршрут из А в В стоимостью, равной 18 усл. ед. (выделе на рис.)
Для сравнения заметим, что другие маршруты стоимости 22, 21, 20, 26 усл. ед.



Пример

Оптимальное планирование замены оборудования.

Проблема своевременной замены устаревшего оборудования одна из насущных проблем горного производства. С течением времени оборудование изнашивается физически и морально, а эксплуатация его на некотором этапе времени становится менее выгодной, чем приобретение и использование нового оборудования.

Пусть в начале планового периода из N лет имеется оборудование возраста t .

Этим оборудованием ежегодно производится продукция стоимостью $\mathbf{r(t)}$ при эксплуатационных расходах для этой цели $\mathbf{u(t)}$.

Остаточная стоимость оборудования $\mathbf{s(t)}$.

В любой год планового периода оборудование можно сохранить или продать по остаточной стоимости и купить новое по цене \mathbf{p} с учетом затрат на эксплуатацию.

Требуется разработать оптимальную политику замены оборудования при условии максимизации прибыли за период времени длительностью N лет.

В соответствии с методикой решения задач динамического программирования начнем процесс оптимизации с конца планового периода, обозначив годы от конца периода к началу: $n = 1, 2, \dots, N$.

Считаем, что к началу последнего года ($n = 1$) у нас имеется оборудование возраста t .

Если оборудование сохранить, то прибыль за последний год составит

$$r(t) - u(t),$$

а если оборудование продать по остаточной стоимости и купить новое, то прибыль к концу последнего года составит

$$s(t) - p + r(0) - u(0),$$

где $r(0)$ — стоимость продукции, произведенной новым оборудованием за год
 $u(0)$ — расходы, связанные с эксплуатацией нового оборудования в течение года

Заменять оборудование выгодно лишь в случае, если

$$s(t) - p + r(0) - u(0) > r(t) - u(t),$$

когда доход от нового оборудования выше, чем от эксплуатации старого.

Обозначим через $f_n(t)$ максимально возможную прибыль за последние n лет планового периода при условии, что в начале периода имеется оборудование возраста t .

В соответствии с этим максимальную прибыль за последний год обозначим через $f_1(t)$ — это наибольшее

$$f_1(t) = \max_t \left\{ \begin{array}{l} r(t) - u(t) \rightarrow \text{сохранение} \\ s(t) - p + r(0) - u(0) \rightarrow \text{замена} \end{array} \right\}.$$

Пусть теперь $n = 2$ (второй этап планирования), т.е. рассматривается период из двух последних лет.

Если к началу этого периода принято решение сохранил оборудование возраста t , то к концу этого периода прибыль, составит $r(t) - u(t)$. За 1 год оборудование постареет и к началу последнего года будет иметь возраст $(t + 1)$ лет. На последнем году от этого оборудования будет получена дополнительная прибыль $f_1(t + 1)$, а общая прибыль за 2 года составит

$$r(t) - u(t) + f_1(t + 1).$$

Если в начале второго года принято решение обновить оборудование, затраты, связанные с продажей старого и приобретением нового оборудования, составят $s(t) - p$, а прибыль от эксплуатации нового оборудования за первый год равна $r(0) - u(0)$.

К концу первого года оборудование постареет и будет иметь возраст 1 год, а прибыль в последнем году составит $f_1(1)$.

Общая прибыль за два года при замене оборудования составит

$$s(t) - p + r(0) - u(0) + f_1(1).$$

Оптимальной в последние два года будет политика, обеспечивающая этот период максимальной прибыль, т.е. равную наибольшему из

$$f_2(t) = \max_t \left\{ \begin{array}{l} r(t) - u(t) + f_1(t+1) \rightarrow \text{сохранение} \\ s(t) - p + r(0) - u(0) + f_1(1) \rightarrow \text{замена} \end{array} \right\}.$$

В общем случае рекуррентное соотношение (уравнение Беллмана) имеет вид

$$f_n(t) = \max_{t,n} \left\{ \begin{array}{l} r(t) - u(t) + f_{n-1}(t+1) \rightarrow \text{сохранение} \\ s(t) - p + r(0) - u(0) + f_{n-1}(1) \rightarrow \text{замена} \end{array} \right\},$$

$$n = 2, 3, \dots; t = 0, 1, 2, \dots$$

Выражение позволяет сформулировать оптимальную политику в области замен устаревшего оборудования с конца планируемого периода, находя

$$f_1(t), f_2(t), \dots, f_N(t).$$

Пример

Проиллюстрируем задачу условным числовым примером. Пусть требуется на пятилетний период $N = 5$ разработать политику замен оборудования не старше $t=10$ лет. Стоимость $r(t)$ продукции и расходы $u(t)$, связанные с эксплуатацией оборудования, даны ниже.

t	0	1	2	3	4	5	6	7	8	9	10
$r(t)$	30	30	29	28	28	26	24	23	22	21	20
$u(t)$	10	11	11	12	12	14	15	17	17	18	20

Ликвидационная стоимость оборудования не зависит равна 2 усл. ед., а цена единицы нового оборудования не меняется и равна 14 усл. ед.

Функциональные уравнение

$$f_1(t) = \max_t \left\{ \begin{array}{l} r(t) - u(t) \rightarrow \text{сохранение} \\ 2 - 14 + 30 - 10 = 8 \rightarrow \text{замена} \end{array} \right\},$$

$$f_n(t) = \max_{t,n} \left\{ \begin{array}{l} r(t) - u(t) + f_{n-1}(t+1) \rightarrow \text{сохранение} \\ 8 + f_{n-1}(1) \rightarrow \text{замена} \end{array} \right\}.$$

Пусть $t = 0$, тогда

$$f_1(0) = \max_t \left\{ \frac{r(0) - u(0)}{8} \right\} = \max_t \left\{ \frac{30 - 10 = 20}{8} \right\} = 20.$$

Так как $20 > 8$, то принимается решение сохранить оборудование. Далее при $t = 1$ находим :

$$f_1(1) = \max_t \left\{ \frac{30 - 11}{8} \right\} = 19.$$

Вывод — при $t = 1$ следует сохранить оборудование. Продолжая аналогично:

$$f_1(6) = \max_t \left\{ \frac{24 - 15}{8} \right\} = 9,$$

т.е. по-прежнему следует вывод о сохранении оборудования. При $t = 7$

$$f_1(7) = \max_t \left\{ \frac{23 - 17}{8} \right\} = 8,$$

что свидетельствует о необходимости замены оборудования при ($t = 7, 6 < 8$).
Далее при $t > 7$ максимизация $f_1(t)$ требует замены оборудования.

Вторую строку таблицы (для двухлетнего периода) заполняем по формуле

$$f_2(t) = \max_i \left\{ \begin{array}{l} r(t) - u(t) + f_1(t+1) \\ 8 + f_1(1) \end{array} \right\}$$
$$f_2(t) = \max_i \left\{ \begin{array}{l} r(t) - u(t) + f_1(t+1) \\ 27 \end{array} \right\}.$$

Придавая значения $t = 0, 1, \dots, 10$, последовательно находим

$$f_2(0) = 39, f_2(1) = 37, f_2(2) = 34 \dots$$

При этом значения $u(t)$, $r(t)$ берем из условия примера, а $f_1(t+1)$

— из первой строки табл.

До $t = 4$ ($n = 2$) убеждаемся, таким образом, что оптимальная политика состоит в сохранении оборудования, а при $t > 5$ лучше произвести его замену.

Третья строка табл. заполняется на основе формулы (для трехлетнего периода $n=3$)

$$f_3(t) = \max_t \left\{ \begin{array}{l} r(t) - u(t) + f_2(t+1) \\ 8 + 37 = 45 \end{array} \right\}.$$

Для $n = 3$ $f_3(0) = 57, f_3(1) = 53, f_3(2) = 50$

необходимо, согласно принципу оптимальности, сохранить оборудование, а уже при $t = 3$ имеем

$$r(3) - u(3) + f_2(4) = 44 < 8 + f_2(1) = 45,$$

и, следовательно, при $t > 3$ следует произвести замену оборудования.

Продолжая вычисления и сравнение, вышеуказанным способом постепенно заполняем табл.

Анализ данных табл. с целью выработки оптимальной политики замены оборудования проводится следующим образом. Область «сохранения» и «замен» оборудования отделим жирной чертой.

Допустим, что в начале пятилетнего периода ($n = 5$) имеется оборудование двухлетнего возраста ($t = 2$) и требуется выработать оптимальное решение относительно этого оборудования на предстоящие пять лет.

В табл. на пересечении $f_5(t)$ и $t = 2$ находим значение прибыли от эксплуатации этого оборудования $f_5(2) = 79$ и замечаем, что это значение находится в области принятия решения сохранения. Это означает, что для достижения в плановом периоде $N = 5$ лет на первом году оборудование двухлетнего возраста надо сохранить.

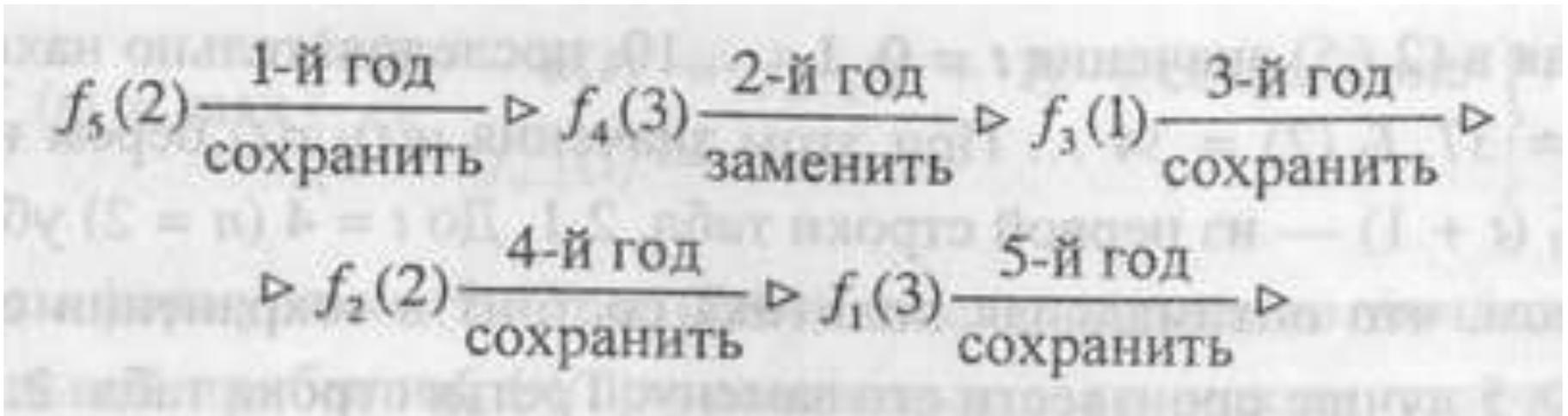
Это оборудование к концу года постареет на один год ($t = 3$), и, следовательно, при $t = 4$ прибыль от эксплуатации его будет меньше, чем в случае замены ($f_4(3) = 61$).

Заменив это оборудование на втором году эксплуатации, при $n = 3$ имеем новое оборудование и через один год работы получим прибыль $f_3(1) = 53$.

Далее, при $n = 4$ приобретенное оборудование стареет на год, следовательно, имеем прибыль $f_2(2) = 34$.

При $n = 5$, точно так же рассуждая находим $f_1(3) = 16$, и все это в области принятия решения сохранения оборудования.

Таким образом, символически формирование оптимальной можно изобразить следующим образом:



Анализируя полученный результат, замечаем, что вместо поиска решения сразу для всех пяти лет находим серию взаимно связанных решений для отрезков времени меньшей длины (1 год), что упростило процедуру планирования.

Для иллюстрации укажем оптимальные решения для оборудования, имеющего на начальный период возраст $t = 0$ (вновь приобретенное),

$$t = 1, t = 10:$$

$$f_5(0) \rightarrow f_4(1) \rightarrow f_3(2) \rightarrow f_2(3) \rightarrow f_1(4) \rightarrow$$

замен не планировать

$$f_5(1) \rightarrow f_4(2) \rightarrow f_3(3) \rightarrow f_2(1) \rightarrow f_1(2)$$

замена

$$f_5(10) \rightarrow f_4(1) \rightarrow f_3(2) \rightarrow f_2(3) \rightarrow f_1(4)$$

- замена

При составлении оптимального плана методом динамического программирования замечаем, что, несмотря на то что в начале планируемого периода имелось разное по возрасту оборудование, к концу планируемого периода характеристики оборудования по возрасту в значительной степени сравнялись, уменьшив количество устаревшего оборудования.

Этим подтверждается важная особенность метода динамического программирования, состоящая в том, что прошлое оптимизируемого процесса не имеет значения при определении принятого решения, а играет роль лишь состояние системы при входе в данный этап.