Глава 4. Динамическое программирование

Если задача математического программирования допускает разбиение процесса её решения на отдельные этапы (шаги), то задача относится к динамическому программированию. Задачи динамического программирования, как и любые другие задачи оптимизации, являются математическими моделями некоторых практических задач принятия решения.

4.1. Метод динамического программирования

Одним из методов решения задач динамического программирования является метод динамического программирования. При помощи метода динамического программирования решаются многошаговые задачи оптимизации. Этот метод применяется в том случае, когда для оптимизации общего критерия качества системы надо многократно выбирать решение, причём система такова, что в ней можно выделить отдельные шаги (этапы), а решения на последующих шагах не оказывают влияния на величину показателя качества, достигнутую на предыдущих шагах. Метод динамического программирования даёт хорошие результаты при малом числе возможных решений на каждом шаге, и прежде всего, когда влияние этих решений описывается лишь небольшим числом переменных. При этих условиях сложность решения задач даже при значительном увеличении числа шагов возрастает лишь слегка.

Обратимся теперь к многошаговым задачам оптимизации, то есть задачам, оптимизацию в которых можно представить в виде ряда последовательных этапов.

Рассмотрим некоторый процесс. Предположим, что состояние некоторого процесса или объекта описывается *n*-мерным вектором $x=(x_1,x_2,...,x_n)$, или, что то же самое, точкой x пространства E_n, которое называют фазовым пространством.

Будем считать, что процесс является *N*-шаговым, то есть он происходит в *N* этапов (шагов) в соответствии со следующей схемой:

$$\underset{\text{начальное состояние}}{x^{(0)}} \to \ldots \to \underbrace{x^{(k-1)} \to x^{(k)}}_{k\text{-й шаг}} \to \ldots \to \underset{\text{конечное состояние}}{x^{(N)}}.$$

Переход между состояниями на *k*-м шаге происходит в соответствии с уравнением состояний:

$$x^{(k)} = f^{(k)}(x^{(k-1)},u^{(k)}), \tag{4.1}$$

где $u^{(k)} \in E_m$ есть *m*-мерный вектор управления, выбираемый на *k*-м шаге, $f^{(k)}(x,u)$-заданная *n*-мерная вектор-функция аргументов $x \in E_n, u \in E_m$.

Таким образом, предполагается, что в результате *k*-го шага процесс переходит в состояние $x^{(k)}$, которое определяется только начальным состоянием $x^{(k-1)}$ этого шага и выбранным на нём вектором управления $u^{(k)}$ и не зависит от “предыстории” процесса до *k*-го шага, т.е. от $x^{(0)},...,x^{(k-2)}$ и $u^{(1)},...,u^{(k-1)}$.

Показателем эффективности k-го шага является заданная числовая характеристика (целевая функция этого шага) $J_k = J_k(x^{(k-1)}, u^{(k)}), k = 1,2,..., N.$

Предположим, что эффективность всего процесса в целом характеризуется целевой функцией вида

$$J(\hat{x}, \hat{u}) = \sum_{k=1}^{N} J_k(x^{(k-1)}, u^{(k)}), \tag{4.2}$$

где $\hat{x} = \{x^{(0)}, x^{(1)},..., x^{(N)}\}$ — набор состояний, называемый фазовой траекторией процесса, а $\hat{u}$ — набор векторов управления, который называется управлением процессом, $\hat{u} = \{u^{(1)}, u^{(2)},..., u^{(N)}\}$.

Таким образом, рассматриваются только аддитивные целевые функции J, представимые в виде суммы целевых функций шагов J_k.

Предположим далее, что на фазовую траекторию и выбор управлений наложены ограничения

$$x^{(k)} \in X_k, k = 1,..., N-1, \tag{4.3}$$

$$u^{(k)} \in U_k(x^{(k-1)}), k = 1,..., N, \tag{4.4}$$

где X_k и $U_k(x^{(k-1)})$ — заданные множества в пространствах E_n и E_m соответственно, причём множество U_k зависит от начального состояния $x^{(k-1)}$ k-го шага.

Ограничения на начальное и конечное состояния процесса $x^{(0)} \in X_0, x^{(N)} \in X_N$ называются начальными и конечными условиями.

При этом множества X_0 и $X_N \subset E_n$ во многих случаях содержат по одной точке (начало и конец фазовой траектории).

Пусть $\hat{u} = \{u^{(1)}, u^{(2)},..., u^{(N)}\}$ — управление процессом, удовлетворяющее ограничениям (4.4) и переводящее его из некоторого начального состояния $x^{(0)} \in X_0$ в некоторое конечное состояние $x^{(N)} \in X_N$ в соответствии с уравнениями (4.1) с учётом ограничений (4.3). Обозначим множество всех таких управлений буквой U.

Многошаговая задача оптимизации формулируется следующим образом: среди всех управлений $\hat{u} \in U$ выбрать такое $\hat{u}^* = \{u^{(1)*}, u^{(2)*},.., u^{(N)*}\}$, для которого целевая функция (4.2) принимает минимальное или максимальное (в зависимости от смысла задачи) значение.

Управление $\hat{u}^*$ и соответствующая ему фазовая траектория $\hat{x}^*$ называются оптимальными.

Условие многошаговой задачи оптимизации будем записывать следующим образом:

$$J(\hat{x}, \hat{u}) = \sum_{k=1}^{N} J_k(x^{(k-1)}, u^{(k)}) \to extr, \tag{4.5}$$

$$x^{(k)} = f^{(k)}(x^{(k-1)}, u^{(k)}), k = 1,..., N, \tag{4.6}$$

$$x^{(k)} \in X_k, k = 1,..., N-1, \tag{4.7}$$

$$u^{(k)} \in U_k(x^{(k-1)}), k = 1,...,N, \qquad (4.8)$$

$$x^{(0)} \in X_0, x^{(N)} \in X_N. \qquad (4.9)$$

Символом extr обозначим минимум или максимум соответствующей функции, в зависимости от смысла задачи.

К многошаговым задачам оптимизации сводятся многие прикладные задачи. Прежде чем решать многие такие задачи, необходимо научиться составлять условия. Для этого рассмотрим на примерах условия задач.

Пример. Сформулировать следующую задачу в виде многошаговой задачи оптимизации (4.5)-(4.9).

С помощью N-ступенчатой ракеты с заданной стартовой массой M в космос запускается межпланетная станция массой m.За время работы каждой ступени ракета получает добавочную скорость $\Delta V = F(y,z)$, *где y — масса, разгоняемая этой ступенью, z — масса самой ступени. Найти такое распределение общей массы M ракеты между её ступенями, при котором конечная скорость станции будет максимальной.*

Р е ш е н и е. Обозначим $u^{(k)}, k = 1,...,N,$ массу *k*-й ступени, считая от межпланетной станции (т.е. на старте работает ступень массой $u^{(N)}$, а в конце разгона – ступень массой $u^{(1)}$). Массу станции вместе с примыкающими к ней *k* ступенями ракеты обозначим $x^{(k)}, k = 0,1,...,N.$

Тогда, очевидно,
$$x^{(k)} = x^{(k-1)} + u^{(k)}, k = 1,...,N;$$
$$x^{(0)} = m, x^{(N)} = M + m.$$

Из условия задачи вытекают следующие ограничения на массы $x^{(k)}$ и $u^{(k)}$: $m \le x^{(k)} \le M + m, 0 \le u^{(k)} \le M$, т.е.
$$x^{(k)} \in X_k = [m; M + m], k = 1,...,N - 1,$$
$$u^{(k)} \in U_k = [0; M], k = 1,...,N.$$

В результате работы *k*-й ступени ракеты приращение скорости станции составит $F(x^{(k-1)}, u^{(k)})$, поэтому её конечная скорость будет равна $\sum_{k=1}^{N} F(x^{(k-1)}, u^{(k)}).$

Таким образом, рассматриваемую задачу можно сформулировать как многошаговую задачу оптимизации следующего вида:
$$J(\hat{x}, \hat{u}) = \sum_{k=1}^{N} F(x^{(k-1)}, u^{(k)}) \to \max,$$
$$x^{(k)} = x^{(k-1)} + u^{(k)}, k = 1,...,N,$$
$$x^{(k)} \in [m; m + M], k = 1,...,N - 1,$$
$$u^{(k)} \in [0; M], k = 1,...,N,$$
$$x^{(0)} = m, x^{(N)} = M + m.$$

Для решения многошаговой задачи оптимизации (4.5) - (4.9) используется метод динамического программирования, основанный на принципе оптимальности Беллмана.

4.2. Принцип оптимальности Беллмана

Оптимальное решение обладает тем свойством, что каковы бы ни были начальное состояние и начальное решение, последующее решение должно быть оптимальным по отношению к состоянию, получающемуся в результате начального решения. Иными словами, принцип оптимальности утверждает, что если в данный момент выбрано не наилучшее решение, то последствия этого нельзя исправить в будущем.

Одной из типичных задач динамического программирования является ***задача распределения ресурсов.***

Задача. *Имеются два предприятия P и Q. При выделении каждому из них на год x единиц средств предприятие P обеспечивает доход 5x единиц и остаток от выделенных средств 0,3x единиц, а предприятие Q — доход 4x единиц и остаток выделенных средств 0,5x единиц.*

Обоим предприятиям выделено на 4 года a=1000 единиц средств. Как нужно распределять их ежегодно между предприятиями, чтобы общий доход за указанный период был максимальным?

Р е ш е н и е. Разобьём весь период продолжительностью в 4 года на 4 этапа, приняв каждый год за один этап. Будем нумеровать этапы, начиная с первого года, и обозначим через x_k, y_k средства, выделяемые соответственно предприятиям P и Q на *k*-м этапе; сумма $x_k + y_k = a_k$ представляет собой общее количество средств, используемых на *k*-м этапе или оставшихся в конце предыдущего (*k*-1)-го этапа. Очевидно, $a_1 = a = 1000$. Доход, получаемый на *k*-м этапе, равен $5x_k + 4y_k$.. Если обозначить через $f_k(a_k)$ максимальный доход, полученный на последних этапах, начиная с *k*-го, т.е. с распределения a_k средств, то функциональное уравнение Беллмана, выражающее принцип оптимальности, будет иметь вид:

$$f_k(a_k) = \max_{0 \le x_k \le a_k} \left[5x_k + 4y_k + f_{k+1}(a_{k+1})\right] \quad (1)$$

Так как $x_k + y_k = a_k$, то для каждого этапа надо выбрать значение величины x_k, при этом $y_k = a_k - x_k$.

Доход на *k*-м этапе составит:

$$5x_k + 4y_k = 5x_k + 4(a_k - x_k) = 4a_k + x_k.$$

Количество средств, используемых на k-м этапе, выразится рекуррентной формулой:

$$a_k = 0{,}3x_{k-1} + 0{,}5y_{k-1} = 0{,}3x_{k-1} + 0{,}5(a_{k-1} - x_{k-1})$$

или

$$a_k = 0{,}1 \cdot (5a_{k-1} - 2x_{k-1}), \quad (2)$$

а функциональное уравнение Беллмана (1) примет вид:

$$f_k(a_k) = \max_{0 \le x_k \le a_k} \left[4a_k + x_k + f_{k+1}(a_{k+1})\right] \quad (3)$$

Полагая последовательно k= 4, 3, 2, 1 (планирование начинается с последнего этапа), получим, используя формулы (2) и (3) и учитывая, что $f_5(a_5)=0$; $k=4,\ f_4(a_4)=\max\limits_{0\le x_4\le a_4}\left[4a_4+x_4\right]=5a_4,$

при $x_4=a_4$ максимум линейной функции $4a_4+x_4$ достигается в точке $x_4=a_4$ — конце отрезка $\left[0;a_4\right], y_4=a_4-x_4=0, a_4=0{,}1\cdot(5a_3-2x_3)$;

$$k=3,\ f_3(a_3)=\max_{0\le x_3\le a_3}\left[4a_3+x_3+5\cdot 0{,}1\cdot(5a_3-2x_3)\right]=\max_{0\le x_3\le a_3}(6{,}5a_3+0x_3)=6{,}5a_3;$$

при любом значении x_3 (функция на отрезке $[0;a_3]$ не зависит от $x_3=const$), $y_3=a_3-x_3; a_3=0{,}1\cdot(5a_2-2x_2)$; $k=2$,

$$f_2(a_2)=\max_{0\le x_2\le a_2}\left[4a_2+x_2+6{,}5\cdot 0.1\cdot(5a_2-2x_2)\right]=\max_{0\le x_2\le a_2}(7{,}25a_2-0{,}3x_2)=7{,}25a_2;$$

при $x_2=0$ (максимум линейной функции $7{,}25a_2-0{,}3x_2$) достигается в точке $x_2=0$ - конце отрезка $\left[0;a_2\right], y_2=a_2-x_2=a_2, a_2=0{,}1\cdot(5a_1-2x_1)$; $k=1$,

$$f_1(a_1)=\max_{0\le x_1\le a_1}\left[4a_1+x_1+7{,}25\cdot 0{,}1\cdot(5a_1-2x_1)\right]=\max_{0\le x_1\le a_1}(7{,}625a_1-0{,}45x_1)=7{,}625a_1,$$

при $x_1=0$ (максимум линейной функции $7{,}625a_1-0{,}45x_1$) достигается в точке $x_1=0$- в конце отрезка $[0;a_1]$), $y_1=a_1-x_1=a_1=a=1000$.

Процесс закончен.

Максимальный доход за 4 года составит $f_1(a_1)=7{,}625\cdot 1000=7625$ единиц. Для получения такого дохода надо в первый и второй годы все средства отдать предприятию Q ($x_1=0, y_1=a_1, x_2=0, y_2=a_2$); на третий год оставшиеся средства можно распределить между предприятиями P и Q произвольным образом ($0\le x_3\le a_3, y_3=a_3-x_3$); и, наконец, на четвёртый год все оставшиеся средства надо передать предприятию P ($x_4=a_4, y_4=0$). Остаток средств, однако, при этом будет зависеть от распределения средств в третьем году. Действительно, $a_1=a=1000$, $a_2=0{,}1\cdot(5a_1-2x_1)=500$, $a_3=0{,}1\cdot(5a_2-2x_2)=250$, $a_4=0{,}1\cdot(5a_3-2x_3)=125-0{,}2x_3$; остаток же после четырёх лет работы будет $a_5=0{,}3a_4=37{,}5-0{,}06x_3$. Величины a_4 и a_5 зависят от x_3. Например, если в третий год все оставшиеся средства передать предприятию P ($x_3=a_3=250, y_3=0$), то $a_5=37{,}5-0{,}06\cdot 250=22{,}5$ единиц, если же предприятию Q ($x_3=0, y_3=250$), то $a_5=37{,}5$ единиц, а если поделить средства пополам ($x_3=y_3=125$), то $a_5=37{,}5-0{,}06\cdot 125=30$ единиц. Вообще в зависимости от x_3 величина a_5 будет меняться в пределах $22{,}5\le a_5\le 37{,}5$. Но в условии задачи не обусловлено никаких требований к остатку. Речь идёт только о максимальной прибыли, которая, независимо от остатка, останется равной 7625 единиц.

Мы рассмотрели задачу, используя первоначальные сведения о целевой функции.

Теперь рассмотрим задачу в общем виде. Оптимальная траектория в задаче (4.5)-(4.9) обладает тем свойством, что любая её завершающая часть, начинающаяся с k-го шага, $k=1,...,N-1$, является оптимальной для остающихся шагов процесса.

Опишем метод динамического программирования.

Заметим, что в формулировке многошаговой задачи оптимизации (4.5)-(4.9) ограничения на фазовую траекторию (4.7) и на конечное состояние (4.9) можно включить в ограничения на выбор управлений, заменив соотношения (4.7) и (4.8) следующим эквивалентным ограничением:

$$u^{(k)} \in \widetilde{U}_k(x^{(k-1)}) = \left\{u^{(k)} \in U_k \middle| f^{(k)}(x^{(k-1)}, u^{(k)}) \in X_k\right\}, k = 1,...,N. \tag{4.10}$$

С учётом этого перепишем формулировку задачи (4.5)-(4.9):

$$J(\hat{x};\hat{u}) = \sum_{k=1}^{N} J_k(x^{(k-1)}, u^{(k)}) \to extr, \tag{4.11}$$

$$x^{(k)} = f^{(k)}(x^{(k-1)}, u^{(k)}), k = 1,...,N, \tag{4.12}$$

$$u^{(k)} \in \widetilde{U}_k(x^{(k-1)}), k = 1,...,N, \tag{4.13}$$

$$x^{(0)} \in X_0. \tag{4.14}$$

Предположим, что в результате начальных $k-1$ шагов процесс перешёл в состояние $x^{(k-1)}$. Рассмотрим задачу оптимизации оставшихся $N-k+1$ шагов, аналогично задаче (4.11)-(4.14). Пусть оптимальное управление $\hat{u}^*(k) = \left\{u^{(k)*},...,u^{(N)*}\right\}$ последних $N-k+1$ шагов и оптимальная траектория этих шагов $\hat{x}^*(k) = \left\{x^{(k-1)}, x^{(k)*},...,x^{(N)*}\right\}$, начатая из состояния $x^{(k-1)}$, найдены.

Целевая функция $J^{(k)}(\hat{x}(k),\hat{u}(k)) = \sum_{i=k}^{N} J_i(x^{(k-1)}, u^{(k)})$ последних $N-k+1$ шагов при $\hat{x}(k) = \hat{x}^*(k), \hat{u}(k) = \hat{u}^*(k)$ принимает оптимальное (т.е. минимальное или максимальное) значение, зависящее от начального состояния $x^{(k-1)}$ фазовой траектории этих шагов, т.е.

$$extr J^{(k)}(\hat{x}(k),\hat{u}(k)) = J^{(k)}(\hat{x}^*(k),\hat{u}^*(k)) = B_k(x^{(k-1)}). \tag{4.15}$$

Функция $B_N(x^{(N-1)})$ из (4.15) называется функцией Беллмана последних $N-k+1$ шагов. Очевидно,

$$B_N(x^{(N-1)}) = \underset{u^{(N)} \in \widetilde{U}_N(x^{(N-1)})}{extr} J_N(x^{(N-1)}, u^{(N)}). \tag{4.16}$$

Кроме того, функции Беллмана связаны между собой следующими соотношениями, вытекающими из принципа оптимальности:

$$B_k(x^{(k-1)}) = \underset{u^{(k)} \in \widetilde{U}_k(x^{(k-1)})}{extr} \left\{B_{k+1}\middle| f^{(k)}(x^{(k-1)}, u^{(k)}) + J_k(x^{(k-1)}, u^{(k)})\right\}, \tag{4.17}$$

$$k = 1,..., N-1.$$

Соотношения (4.16) и (4.17), позволяющие последовательно найти функции Беллмана $B_N(x^{(N-1)}),.B_{N-1}(x^{(N-2)}),...,.B_1(x^{(0)})$, называются уравнениями Беллмана.

Находя функции $B_k(x^{(k-1)}), k = N, N-1,...,1$ из (4.16) и (4.17), мы одновременно определяем и управления $u^{(k)*}(x^{(k-1)})$, которым отвечают оптимальные значения соответствующих величин $Z_N = J_N(x^{(N-1)}, u^{(N)})$ и $Z_k = B_{k+1}\left[f^{(k)}(x^{(k-1)}, u^{(k)})\right] + J_k(x^{(k-1)}, u^{(k)})$, $k = N-1, N-2,...,1$ из правых частей (4.16), (4.17):

$$Z_N\left[x^{(N-1)}, u^{(N)*}(x^{(N-1)})\right] = \underset{u^{(N)} \in \widetilde{U}_N(x^{(N-1)})}{extr} Z_N(x^{(N-1)}, u^{(N)}) = B_N(x^{(N-1)}), \tag{4.18}$$

$$Z_k\left[x^{(k-1)},u^{(k)*}(x^{(k-1)})\right]=\underset{u^{(k)}\in\widetilde{U}_k(x^{(k-1)})}{extr} Z_k(x^{(k-1)},u^{(k)})=B_k(x^{(k-1)}), \qquad (4.19)$$

$k=N-1,...,1$.

Управления $u^{(k)*}(x^{(k-1)}), k=1,...,N,$ называются условными оптимальными управлениями, а процесс их нахождения — условной оптимизацией. Отметим, что управления $u^{(k)*}(x^{(k-1)}),$ найденные в соответствии с (4.19), удовлетворяют принципу оптимальности, т.е. в зависимости от начального состояния $x^{(k-1)}$ управление $u^{(k)*}(x^{(k-1)}), k=1,...,N-1,$ учитывает оптимизацию не только k-го шага, но и следующих за ним $N-k$ шагов.

Итак, в результате условной оптимизации находятся функции Беллмана $B_k(x^{(k-1)})$ и условные оптимальные управления $u^{(k)*}(x^{(k-1)}), k=1,...,N.$

После этого можно осуществить безусловную оптимизацию в задаче (4.11) - (4.14), то есть определить искомое оптимальное управление процессом $\hat{u}^{*}=\left\{u^{(1)*},...,u^{(N)*}\right\}$ и оптимальную фазовую траекторию $\hat{x}^{*}=\left\{x^{(0)*},...,x^{(N)*}\right\}$ следующим образом.

Так как функция Беллмана $B_1(x^{(0)})$ для каждого начального состояния $x^{(0)}\in X_0$ равна оптимальному значению целевой функции N шагов, т.е. всего процесса, начатого из состояния $x^{(0)}$, то оптимальное начальное условие $x^{(0)*}\in X_0$ находим из соотношения

$$B_1(x^{(0)*})=\underset{x^{(0)}\in X_0}{extr} B_1(x^{(0)}) \qquad (4.20)$$

(если множество X_0 из (4.14) состоит из единственной точки $x^{(0)}$, то полагаем $x^{(0)*}=x^{(0)}$).

Далее, используя найденные условные оптимальные управления, а также уравнения состояний (4.12), последовательно находим $u^{(1)*},x^{(1)*},u^{(2)*},x^{(2)*},...,u^{(N)*},x^{(N)*}$ из следующих соотношений:

$$\begin{aligned}
&u^{(1)*}=u^{(1)*}(x^{(0)*}), x^{(1)*}=f^{(1)}(x^{(0)*},u^{(1)*}),\\
&u^{(2)*}=u^{(2)*}(x^{(1)*}), x^{(2)*}=f^{(2)}(x^{(1)*},u^{(2)*}),\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&u^{(N)*}=u^{(N)*}(x^{(N-1)*}), x^{(N)*}=f^{(N)}(x^{(N-1)*},u^{(N)*}).
\end{aligned} \qquad (4.21)$$

Теперь рассмотрим алгоритм метода динамического программирования.

Этап I (условная оптимизация).

<u>*Ш а г 1*</u>. Находим условное оптимальное управление $u^{(N)*}(x^{(N-1)})$ и функцию Беллмана $B_N(x^{(N-1)})$ в соответствии с (4.16), (4.18);

<u>*Ш а г 2.*</u> Используя результаты первого шага, находим $u^{(N-1)*}(x^{(N-2)})$ и $B_{N-1}(x^{(N-2)})$ с помощью равенств (4.17), (4.19) при $k=N-1$;

…………………………………………………………………………

<u>*Ш а г N.*</u> Используя результаты (*N-1*)-го шага, определяем $u^{(1)*}(x^{(0)})$ и $B_1(x^{(0)})$ в соответствии с (4.17) и (4.19) при $k=1$.

Этап II (безусловная оптимизация).

<u>*Ш а г 0*</u>. Находим оптимальное начальное состояние $x^{(0)*}$ в соответствии с (4.20);

Ш а г 1. Определяем оптимальное управление $u^{(1)*}$ и конечное состояние $x^{(1)*}$ первого шага процесса по формулам (4.21);

Ш а г 2. Используя результаты предыдущего шага, находим $u^{(2)*}$ и $x^{(2)*}$ в соответствии с (4.21);

…………………………………………………………………………..

Ш а г N. С помощью результатов ($N-1$)-го шага определяем $u^{(N)*}$ и $x^{(N)*}$ по формулам (4.21).

Окончательно имеем $\hat{u}^{*} = \{u^{(1)*},...,u^{(N)*}\}$, $\hat{x}^{*} = \{x^{(0)*},...,x^{(N)*}\}$.

Для многих экономических и производственных задач характерной является дискретная модель, предполагающая, что величины, описывающие процесс, могут принимать только дискретный ряд значений. Функциональные зависимости в таких задачах задаются, как правило, в виде таблиц, а не аналитически. Однако общая схема их решения методом динамического программирования остаётся без изменений.

Рассмотрим решение следующего примера.

Пример. *Общая сумма в 4 млн. руб. распределяется между тремя предприятиями в количествах, кратных 1 млн. руб. В результате выделения средств k-му предприятию в размере* u *оно даёт доход $J_k(u), k=1,2,3,$ величина которого может быть найдена из табл. 1.*

Таблица 1

U	0	1	2	3	4
$J_1(u)$	0	5	9	11	12
$J_2(u)$	0	4	8	12	14
$J_3(u)$	0	7	9	10	11

Распределить средства между предприятиями так, чтобы их суммарный доход был максимальным.

Р е ш е н и е. Обозначив средства, выделенные k-му предприятию ($k=1,2,3$), символом $u^{(k)}$, а сумму средств, выделенных предприятиям с номерами от 1 до k, символом $x^{(k)}$ сформулируем рассматриваемую задачу как многошаговую задачу (4.11) - (4.14):

$$J(\hat{x},\hat{u}) = \sum_{k=1}^{3} J_k(u^{(k)}) \to \max,$$

$$x^{(k)} = x^{(k-1)} + u^{(k)}, k = 1,2,3,$$

$$u^{(k)} \in \left[0; u - x^{(k-1)}\right] \cap Z, k = 1,2,3,$$

$$x^{(0)} = 0, x^{(3)} = 4.$$

Для решения этой задачи применим метод динамического программирования.

Этап I (условная оптимизация).

Ш а г 1.

Найдём $B_3(x^{(2)}) = \max\limits_{u^{(3)} \in \left[0;4-x^{(2)}\right] \cap Z} J_3(u^{(3)}).$

Так как функция $Z_3(x^{(2)},u^{(3)}) = J_3(u^{(3)})$ является возрастающей функцией аргумента $u^{(3)}$ (см. таблицу 1), то её максимум достигается при максимальном допустимом значении $u^{(3)}$, т.е.

$$u^{(3)*}(x^{(2)}) = \left[4 - x^{(2)}\right] ([\ ]\text{-целая часть}) \qquad (*)$$

Отсюда $B_3(x^{(2)}) = Z_3(x^{(2)},u^{(3)*}(x^{(2)})) = J_3(\left[4 - x^{(2)}\right])$. Значения $B_3(x^{(2)})$, найденные с помощью таблицы 1 представлены в табл. 2.

Таблица 2

$x^{(2)}$	0	1	2	3	4
$B_3(x^{(2)})$	11	10	9	7	0

Шаг 2.

Вычислим $B_2(x^{(1)}) = \max\limits_{u^{(2)} \in [0;4-x^{(1)}] \cap Z} \left\{B_3(x^{(1)} + u^{(2)}) + J_2(u^{(2)})\right\}$

Для нахождения максимума функции $Z_2(x^{(1)},u^{(2)}) = B_3(x^{(1)} + u^{(2)}) + J_2(u^{(2)})$ составляем табл. 3 значений этой функции, используя данные табл. 1 и 2.

Таблица 3

$x^{(1)}$	$u^{(2)}$				
	0	1	2	3	4
0	11	14	17	$\langle 19 \rangle$	14
1	10	13	$\langle 15 \rangle$	12	-
2	9	$\langle 11 \rangle$	8	-	-
3	$\langle 7 \rangle$	4	-	-	-
4	$\langle 0 \rangle$	-	-	-	-

В таблице 3 скобками выделены максимальные по $u^{(2)}$ значения функции $Z_2(x^{(1)},u^{(2)})$, соответствующие различным значениям $x^{(1)}$.

Используя табл. 3, находим функции $B_2(x^{(1)})$ и $u^{(2)}(x^{(1)})$, представив их значения в табл. 4 и 3.

Таблица 4

$x^{(1)}$	0	1	2	3	4
$B_2(x^{(1)})$	19	15	11	7	0

Таблица 5

$x^{(1)}$	0	1	2	3	4
$u^{(2)*}(x^{(1)})$	3	2	1	0	0

Таблица 6

$u^{(1)}$	0	1	2	3	4
$Z_2(0,u^{(1)})$	19	$\langle 20\rangle$	$\langle 20\rangle$	19	12

<u>*Ш а г 3.*</u>

Так как множество X_0 состоит из единственной точки $x^{(0)}=0$, то найдём только $B_1(0)$ и $u^{(1)*}(0)$:

$$B_1(0)=\max_{u^{(1)}\in[0;4]\cap Z}\left\{B_2(0+u^{(1)})+J_1(u^{(1)})\right\}$$

Для определения максимума в правой части последнего равенства составим табл. 6 значений функции $Z_2(0,u^{(1)})=B_2(u^{(1)})+J_1(u^{(1)})$, которые найдём с помощью таблиц 1 и 4.

Из табл. 6 видно, что $B_1(0)=20$, причём $u^{(1)*}(0)=1$, или $u^{(1)*}(0)=2$, то есть в данной задаче существует два оптимальных управления и две оптимальные траектории.

Этап II (безусловная оптимизация).

<u>*Ш а г 1.*</u>

a) Пусть $u^{(1)*}=1$. Тогда $x^{(1)*}=x^{(0)*}+u^{(1)*}=1$.

b) Пусть $u^{(1)*}=2$. Тогда $x^{(1)*}=x^{(0)*}+u^{(1)*}=2$.

<u>*Ш а г 2.*</u>

a) Для $x^{(1)*}=1$ имеем $u^{(2)*}=u^{(2)*}(1)=2$ (смотри табл. 5), $x^{(2)*}=x^{(1)*}+u^{(2)*}=3$.

b) Для $x^{(1)*}=2$ получаем $u^{(2)*}=u^{(2)*}(2)=1$ (смотри табл. 5), $x^{(2)*}=x^{(1)*}+u^{(2)*}=3$.

<u>*Ш а г 3.*</u>

Так как для обеих оптимальных фазовых траекторий $x^{(2)*}=3$, то из (*) находим $u^{(3)*}=u^{(3)*}(3)=1$, $x^{(3)*}=x^{(2)*}+u^{(3)*}=4$.

Окончательно получаем $\hat{u}^{*}=\{1,2,1\}$ или $\hat{u}^{*}=\{2,1,1\}$ и соответственно $\hat{x}^{*}=\{0,1,3,4\}$ или $\hat{x}^{*}=\{0,2,3,4\}$.

Таким образом, существует два оптимальных варианта распределения средств предприятиям:

1) первому предприятию выделяется 1 *млн.руб.*, второму — 2 *млн.руб.* и третьему — 1 *млн.руб.*;
2) первому — 2, второму — 1 и третьему — 1 *млн.руб.*

В обоих случаях суммарный доход предприятий составит $B_1(0)=20$ *млн.руб.*

4.3. Решение задач методом динамического программирования

1. Задача об оптимальной загрузке рюкзака

Исследуемый здесь вопрос имеет многочисленные приложения, в частности, к задачам рационального раскроя. Для наглядности мы рассмотрим его первоначально в более красочной интерпретации.

Во многих старинных сказках доброму молодцу в награду за спасение малолетнего царевича предлагают выбрать из царского хранилища любые предметы, которые он сможет унести. Если при этом в хранилище имеются в достаточном количестве предметы m типов, то выбираемый набор характеризуется m-мерным вектором

$$x = (x_1, x_2, ..., x_m) \tag{4.1.1}$$

с целыми неотрицательными компонентами, указывающими количество выбранных предметов каждого типа. Допустим далее, что каждый предмет типа $i = \overline{1,m}$ имеет стоимость c_i и вес a_i, а предельный вес набора характеризуется величиной b.

В указанных условиях, естественно, добрый молодец будет стремиться максимизировать общую стоимость набора

$$\mu(x) = \sum_{i=1}^{m} c_i x_i$$

при единственном линейном ограничении

$$\sum_{i=1}^{m} a_i x_i \le b.$$

Перебор всех допустимых наборов в общем случае, очевидно, исключается. В связи с этим комплектование набора будем трактовать как многошаговый процесс, состоящий из последовательного выбора отдельных предметов.

В соответствии с общими принципами динамического программирования наряду с интересующей нас исходной задачей рассмотрим семейство аналогичных задач

$$f(z) = \max_{x \in X_z} \mu(x), z \in [0, b],$$

где через X_z обозначено множество неотрицательных целочисленных векторов (4.1.1), отвечающих наборам, в которых общий вес предметов не превосходит соответствующей величины z.

Пусть $z_0 = \min_{i=\overline{1,n}} a_i$. Тогда при всех $z \in [0, z_0]$ соответствующие множества X_z, очевидно, состоят из одного нулевого элемента и, следовательно, $f(z) = 0$ для всех таких z. Что касается $z \in [z_0, b]$, то для них, как легко проверить, справедливы следующие рекуррентные соотношения:

$$f(z) = \max_{i \in I_z} [c_i + f(z - a_i)], \tag{4.1.2}$$