

# Задачи динамического программирования.

*Динамическое программирование представляет собой метод оптимизации многошаговых процессов принятия решения, позволяющий указать пути исследования целого класса экстремальных задач.*

Метод оказывается весьма эффективным при анализе задач с аддитивной целевой функцией

$$z = \sum_{j=1}^n f_j(x_j)$$

Предположим, что интересующий нас процесс можно разбить на N шагов, а действия, совершаемые на i-ом шаге, характеризуются совокупностью показателей

$$\vec{X}_i^+ = (x_{i1}, x_{i2}, \dots, x_{im})$$

Состояние процесса к началу этого шага имеет характеристику в виде набора параметров

$$\vec{Q}_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iM})$$

Обычно предпринимаемые действия не являются полностью произвольными, а зависят от состояния

$$\vec{Q}_i = (\pi_{i1}, \pi_{i2}, \dots, \pi_{iM})$$

которое возникло перед  $i$ -ым шагом, т.е.  $\vec{X}_i = \vec{X}_i(\vec{Q}_i)$

Очевидно, результирующее значение критерия  $Z$ , получаемое в конце процесса, будет определяться теми  $X$ , которые были приняты до этого

$$Z = Z(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_N)$$

Как выбрать  $\vec{X}_1, \vec{X}_2, \dots, \vec{X}_N$

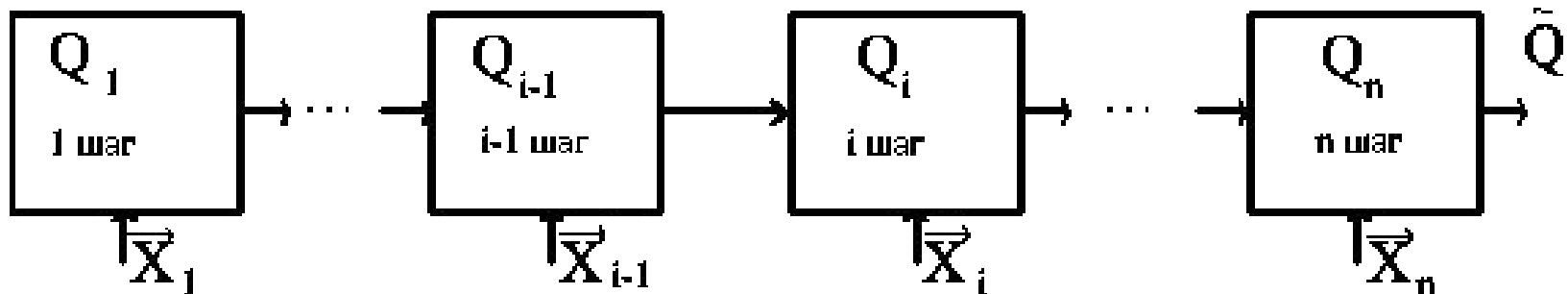
чтобы величина  $Z$  приняла экстремальное значение?

Ответ можно получить, рассматривая  $Z$  как функцию переменных

$$x_i, (i = \overline{1, N}, s = \overline{1, M})$$

и находя экстремум  $z$  одним из известных способов, однако этот путь не всегда прост (особенно при большом числе переменных).

Появляется идея провести оптимизацию поэтапно, анализируя последовательно каждый шаг процесса в поисках наилучших вариантов его продолжения.



Эта идея лежит в основе метода динамического программирования, реализующего принцип последовательной оптимизации. Следовательно, важным условием применимости рассматриваемого метода является возможность разбиения процесса принятия решений на ряд однотипных шагов или этапов, каждый из которых планируется отдельно, но с учетом результатов, полученных на других шагах.

Динамическое программирование основывается на двух важных принципах-оптимальности и вложения.

**Принцип оптимальности формулируется следующим образом:**

Каково бы ни было состояние системы в результате какого-то числа шагов, мы должны выбирать управление на ближайшем шаге так, чтобы оно, в совокупности с оптимальным управлением на всех последующих шагах, приводило к максимальному выигрышу на всех оставшихся шагах, включая данный.

**Принцип вложения утверждает, что природа задачи, допускающей использование метода динамического программирования, не меняется при изменении количества шагов, т. е. форма такой задачи инвариантна относительно  $N$ .** В этом смысле всякий конкретный процесс с заданным  $N$  оказывается как бы вложенным в семейство подобных ему процессов и может рассматриваться с более широких позиций.

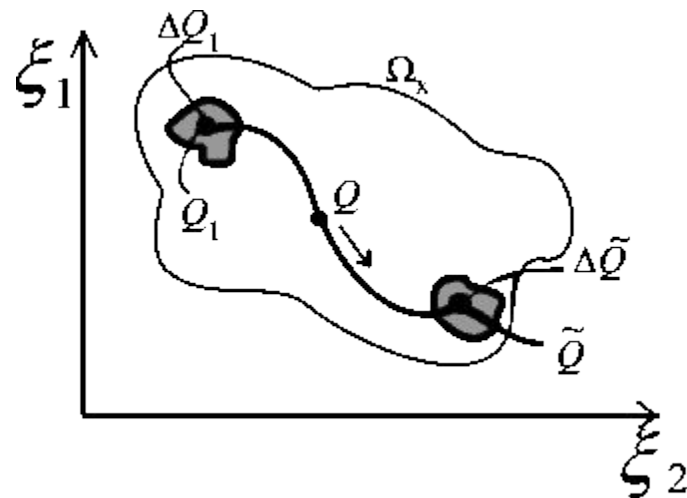
Реализация названных принципов дает гарантию того, что, во-первых, решение, принимаемое на очередном шаге, окажется наилучшим с точки зрения всего процесса (*а не "узких интересов" отдельного этапа*) и, во-вторых, последовательность решений одношаговой, двухшаговой и т. п. задач приведет к решению исходной  $N$ -шаговой задачи.

В общем случае необходимо учитывать некоторые условия, накладываемые на начальное и конечное состояния системы:

$$Q_1 \in \Delta Q_1 \quad \tilde{Q} \in \Delta \tilde{Q}$$

$\Delta Q_1$  область начальных состояний  
 $\Delta \tilde{Q}$  область конечных состояний

Если состояние системы характеризуется двумя фазовыми переменными  $\xi_1, \xi_2$   
 то процесс управления представляет собой перемещение точки из  $Q_1 \in \Delta Q_1$   
 в  $\tilde{Q} \in \Delta \tilde{Q}$



Владелец автомашины эксплуатирует ее в течение  $m$  лет. В начале каждого года он может принять одно из трех решений:

1. *Продать машину и заменить ее новой;*
2. *Ремонтировать ее и продолжать эксплуатацию;*
3. *Продолжать эксплуатацию без ремонта.*

### ***Шаговое управление-выбор одного из этих трех решений.***

Непосредственно числами они не выражаются, но можно приписать первому численное значение 1, второму 2, третьему 3. Какие нужно принять решения по годам (т. е. как чередовать управления 1, 2,3), чтобы суммарные расходы на эксплуатацию, ремонт и приобретение новых машин были минимальны?

Показатель эффективности (в данном случае это не "выигрыш", а "проигрыш", но это неважно) равен

$$Z = \sum_{i=1}^N z_i$$

$z_i$  - расходы в  $i$ -м году

Величину  $Z$  требуется обратить в минимум. Управление операцией в целом представляет собой какую-то комбинацию чисел 1, 2, 3, например:

$$\vec{X} = (3, 3, 2, 2, 2, 1, 3, \dots)$$

Метод **динамического программирования** позволяет с успехом решать экономические задачи. *Рассмотрим одну из простейших таких задач:*  
В нашем распоряжении имеется какой-то запас средств (ресурсов)  $K$ , который должен быть распределен между  $m$  предприятиями

$$P_1, P_2, \dots, P_m$$

Каждое из предприятий  $P_i$  при вложении в него каких-то средств  $x$  приносит доход, зависящий от  $x$ , т. е. представляющий собой какую-то функцию  $\varphi_i(x)$

Все функции  $\varphi_i(x)$  ( $i = 1, 2, \dots, m$ ) . известны

Спрашивается, как нужно распределить средства  $K$  между предприятиями, чтобы в сумме они дали максимальный доход?

Управляемая система  $S$  в данном случае-средства или ресурсы, которые распределяются. Состояние системы  $S$  перед каждым шагом характеризуется одним числом  $Q$  - наличным запасом еще не вложенных средств.

## Решение задачи динамического программирования

Процедура построения оптимального управления методом динамического программирования распадается на две стадии: предварительную и окончательную.

На предварительной стадии для каждого шага определяется условное оптимальное управление, зависящее от состояния  $-Q$  системы (*достигнутого в результате предыдущих шагов*), и условный оптимальный выигрыш на всех оставшихся шагах, начиная с данного, также зависящий от состояния  $Q$ .

На окончательной стадии для каждого шага определяется окончательное (безусловное) оптимальное управление.

**Предварительная (условная) оптимизация** проводится по шагам, в обратном порядке:

*от последнего шага к первому;*

**Окончательная (безусловная) оптимизация**-также по шагам, но в естественном порядке:

*от первого шага к последнему.*



**Введем так называемую функцию Беллмана в соответствии с условием оптимальности**

$$Z_i(Q_i) = \max_{\vec{X}_i} \{z_i(Q_i, \vec{X}_i) + Z_{i+1}(\varphi_i(Q_i, \vec{X}_i))\}$$

$Z_i(Q)$  -условный оптимальный выигрыш, получаемый на всех последующих шагах, начиная с  $i$ -го и до конца;  
он достигается при оптимальном управлении на всех этих шагах и равен максимальному выигрышу, который можно получить на всех этих шагах вместе, если перед их началом система находится в состоянии  $Q_i$

$$\vec{X}_i(Q_i)$$

условное оптимальное управление на  $i$ -ом шаге, которое, совместно с оптимальным управлением на всех последующих шагах, обращает выигрыш на всех шагах, начиная с данного, в максимум.

$$Q_{i+1} = \varphi_i(Q_i, X_i)$$

примененное управление на  $i$ -м шаге приведшее ее в новое состояние

$$z_i = z_i(Q, \vec{X}_i)$$

Выигрыш на текущем шаге

**уравнение Беллмана** позволяет определить функцию  $Z_i(Q)$ , если известна следующая за ней по порядку  $Z_{i+1}(Q)$

Что касается функции  $Z_N(Q)$  (*условный оптимальный выигрыш на последнем шаге*), то она может быть определена очень просто. Действительно, за последним шагом нет следующего, и нужно обратить в максимум выигрыш на этом шаге:

$$Z_N(Q_N) = \max_{\vec{X}_N} \{ z_N(Q_N, X_N) \}$$

Максимум в формуле берется не по всем возможным управлениям  $X_N$  на N-ом шаге, а только по тем, которые приводят в систему в заданную область  $\Delta\tilde{Q}$  т.е. по тем, для которых

$$\varphi_N(Q, \vec{X}_N) \in \Delta\tilde{Q}$$

управление  $\vec{X}_N = \vec{x}_N(Q)$  при котором достигается максимум выигрыша, и есть условное оптимальное управление на последнем шаге.

Зная  $Z_N(Q)$  можно по общей формуле Беллмана

полагая в ней  $i+1=n$ , найти функцию  $Z_{N-1}(Q)$  и соответствующее оптимальное управление  $X_{N-1}(Q)$  и так далее

до последнего шага от конца т.е. первого шага, для которого будут найдены функции  $Z1(Q)$  и  $X1(Q)$  . Функция  $Z1(Q)$  есть условный оптимальный выигрыш за всю операцию, т.е. на всех шагах начиная с первого и до последнего (*если первый шаг начинается с определенного состояния  $Q1$* ).

**Таким образом предварительная оптимизация закончена-**  
найжены условный оптимальный выигрыш и условное оптимальное управление для каждого шага.

**Рассмотрим вторую стадию оптимизации**-нахождение безусловного или окончательного оптимального управления

•Начнем с первого шага.

$$x = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_N)$$

Предположим, что исходное состояние  $Q_1$  нам полностью известно. Подставим это состояние  $Q_1$  в формулу для условного оптимального выигрыша  $Z_1(Q)$ .

Получим:

$$Z_{\max} = Z_1(Q)$$

Одновременно найдем оптимальное управление на первом шаге  $\vec{x}_1 = \vec{x}_1(Q)$  Далее, зная исходное состояние  $Q_1$  и управление  $X_1$ , можем найти состояние системы после первого шага:  $Q_2^* = Q_2(Q_1)$  Зная это состояние можно найти оптимальное управление на втором шаге  $\vec{x}_2 = \vec{x}_2(Q_2^*)$  затем и т. д. Таким образом, идя по цепочке

$$Q_1 \rightarrow x_1(Q_1) \rightarrow (Q_2^*) \rightarrow x_2(Q_2^*) \rightarrow \dots \rightarrow Q_N^* \rightarrow x_N(Q_N^*) \rightarrow \tilde{Q}$$