Лекция 3

1. *Понятие о динамическом программировании*
2. *Принцип оптимальности и уравнения Беллмана*
3. *Задача о выборе оптимального пути и ее решение*
4. *Задача о распределении средств между двумя предприятиями*
5. *Решение задачи методом динамического программирования.*

1. Понятие о динамическом программировании.

Динамическое программирование – метод оптимизации, приспособленный к ***операциям,*** в которых процесс принятия решения может быть разбит на этапы (шаги). (Р.Беллман (1920) – американский математик).

Операция – управляемое мероприятие, направленное на достижение некоторой цели

Рассматривается некоторый ***управляемый процесс***. В результате ***управления*** система (***объект управления***) ***S*** переводится из начального состояния S_0 в конечное состояние S^*. Управление можно разбить на ***n*** шагов, то есть решение принимается последовательно на каждом шаге.

$$S_0 \xrightarrow{y_1} S_1 \xrightarrow{y_2} S_2 \xrightarrow{y_3} \dots \xrightarrow{y_{n-1}} S_{n-1} \xrightarrow{y_n} S^*$$

Пусть y_k- управление на ***k***-м шаге (***k=1,2,…,n;*** y_k- число, точка в **n-**мерном пространстве, функция, качественный признак и т.д.).

Пусть $Y=(y_1,y_2,\dots,y_n)$ – управление, переводящее систему из состояния S_0 в состояние S^*.

Показатель эффективности – целевая функция, зависит от начального состояния и управления

$$Z = F(S_0, Y)$$

Основные предположения:

1.Состояние $\boldsymbol{S_k}$ системы в конце $\boldsymbol{k}$-го шага зависит только от предшествующего состояния $\boldsymbol{S_{k-1}}$ и управления на $\boldsymbol{k}$-м шаге $\boldsymbol{y_k}$ ***(отсутствие последействия).***

$$S_k = \Phi\left(S_{k-1}, y_k\right)$$

2.Целевая функция является ***аддитивной*** от показателя эффективности каждого шага, т.е. если показатель эффективности $\boldsymbol{k}$-го шага равен

$$Z_k = f_k\left(S_{k-1}, y_k\right) \qquad k = 1, 2, \ldots, n$$

то

$$Z = F\left(S_0, Y\right) = \sum_{k=1}^{n} f_k\left(S_{k-1}, y_k\right)$$

Задача. Определить такое допустимое управление $\boldsymbol{Y}$, переводящее систему $\boldsymbol{S}$ из состояния $\boldsymbol{S_0}$ в состояние $\boldsymbol{S^*}$, при котором целевая функция принимает наибольшее (наименьшее) значение.

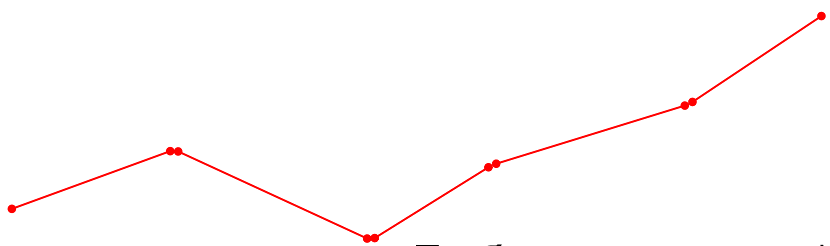
Такое управление называют ***оптимальным.***

2.Принцип оптимальности и уравнения Беллмана

В 1953 году Р.Беллманом был сформулирован принцип:

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

Оптимальная траектория



Любая часть этой ломаной будет являться оптимальной траекторией относительно начала и конца

На каждом шаге решение Y_k нужно выбирать «с оглядкой», так как этот выбор влияет на последующее состояние S_k и дальнейший процесс управления, зависящий от S_k. ***Но есть один шаг, последний, который можно для любого состояния*** S_{n-1} ***планировать локально-оптимально, исходя только из соображений этого шага.***

Пусть $Z_n^*(S_{n-1}) = \varphi_n(S_{n-1})$ - максимум целевой функции (показателя эффективности) n-го шага при условии, что к началу последнего шага

система ***S*** была в произвольном состоянии $\boldsymbol{S_{n-1}}$, а на последнем шаге управление было оптимальным.

$Z_n^*(S_{n-1})$ называется условным максимумом целевой функции на ***n*** –м шаге.

$$Z_n^*(S_{n-1}) = \varphi_n(S_{n-1}) = \max_{\{y_n\}} f_n(S_{n-1}, y_n)$$

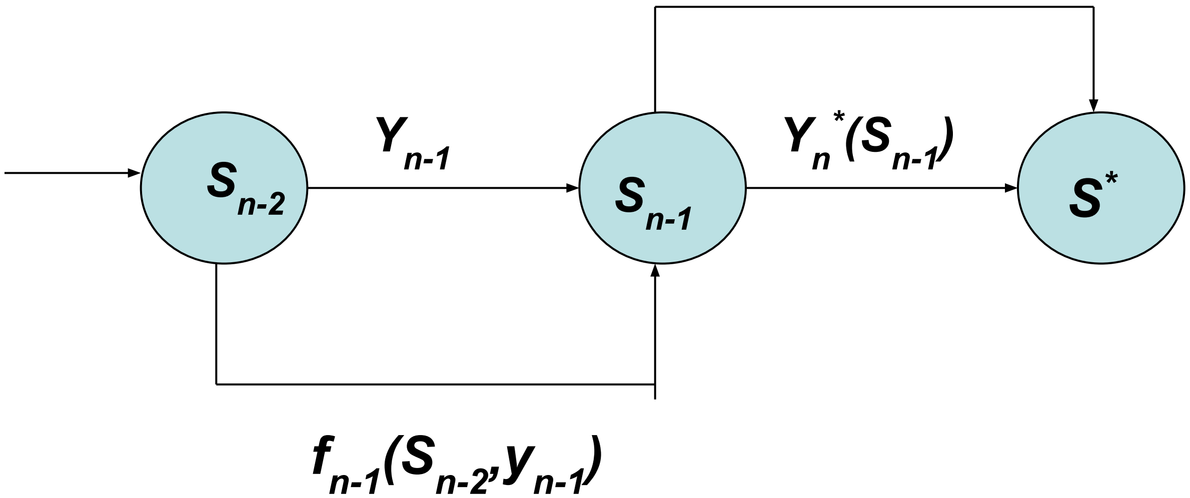
Максимизация ведется по всем допустимым управлениям $\boldsymbol{y_n}$

Управление $\boldsymbol{y_n}$, при котором достигается $Z_n^*(S_{n-1})$, также зависит от $\boldsymbol{S_{n-1}}$

и называется условным оптимальным управлением на ***n*** -м шаге. Оно обозначается через $\boldsymbol{y_n^*(S_{n-1})}$

Рассмотрим два последних шага (двухшаговая задача)

$$Z_n^*\left(S_{n-1}\right)=\varphi_n\left(S_{n-1}\right)$$ *условный оптимальный выигрыш на* ***n-м*** *шаге*



$f_{n-1}\left(S_{n-2},y_{n-1}\right)$ - значение целевой функции ***n-1*** –го шага при произвольном

управлении $\boldsymbol{y_{n-1}}$ и состоянии $\boldsymbol{S_{n-2}}$. Значение целевой функции на двух последних шагах равно

$$f_{n-1}\left(S_{n-2},y_{n-1}\right)+\varphi_n\left(S_{n-1}\right)$$

Тогда $Z_{n-1}^*\left(S_{n-2}\right)=\varphi_{n-1}\left(S_{n-2}\right)$ называется условным максимумом целевой функци

при оптимальном управлении на двух последних шагах

Соответствующее управление y_{n-1} на **(n-1)**-м шаге обозначается через

$y_{n-1}^{*}(S_{n-2})$ *и* называется условным оптимальным управлением на **(n-1)**-м шаге

$$Z_{n-1}^{*}\left(S_{n-2}\right)=\varphi_{n-1}\left(S_{n-2}\right)=\max_{\{y_{n-1}\}}\left\{f_{n-1}\left(S_{n-2},y_{n-1}\right)+\varphi_{n}\left(S_{n-1}\right)\right\}$$

где $$S_{n-1}=\Phi\left(S_{n-2},y_{n-1}\right)$$

Уравнения Беллмана имеют вид

$$\boxed{Z_{n}^{*}\left(S_{n-1}\right)=\varphi_{n}\left(S_{n-1}\right)=\max_{\{y_{n}\}}\left\{f_{n}\left(S_{n-1},y_{n}\right)\right\}}$$

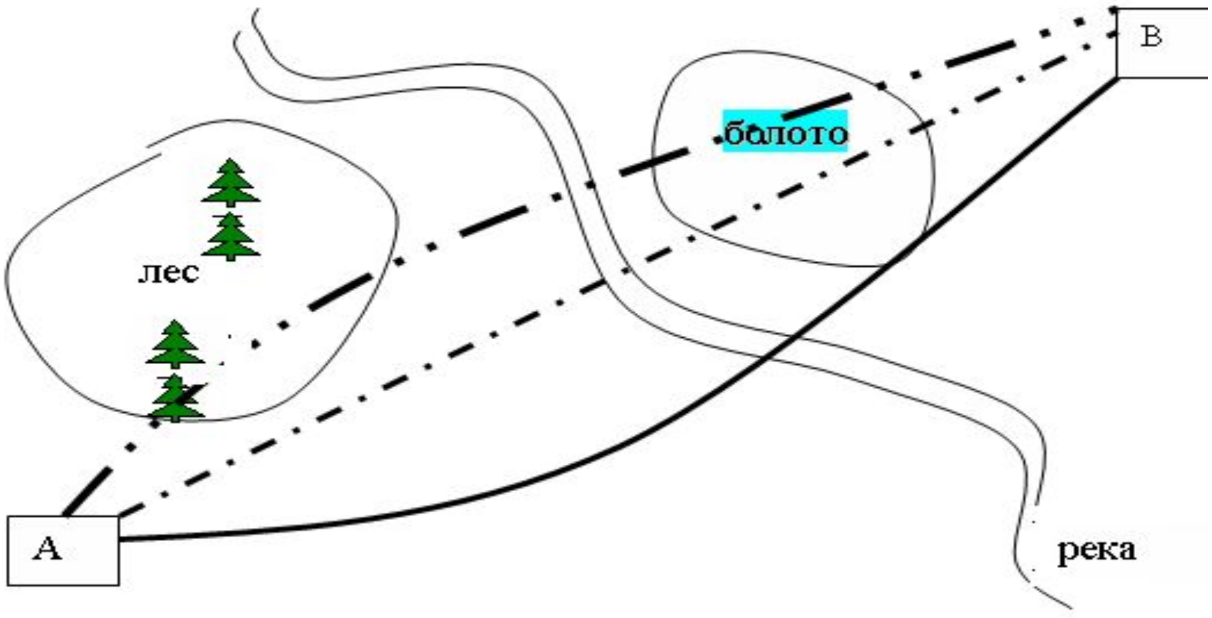
$$\boxed{Z_{k}^{*}\left(S_{k-1}\right)=\varphi_{k}\left(S_{k-1}\right)=\max_{\{y_{k}\}}\left\{f_{k}\left(S_{k-1},y_{k}\right)+\varphi_{k+1}\left(S_{k}\right)\right\}}\quad k=n-1,\ n-2,\dots,2,\ 1$$

$$S_{k}=\Phi\left(S_{k-1},y_{k}\right)$$ уравнение состояния

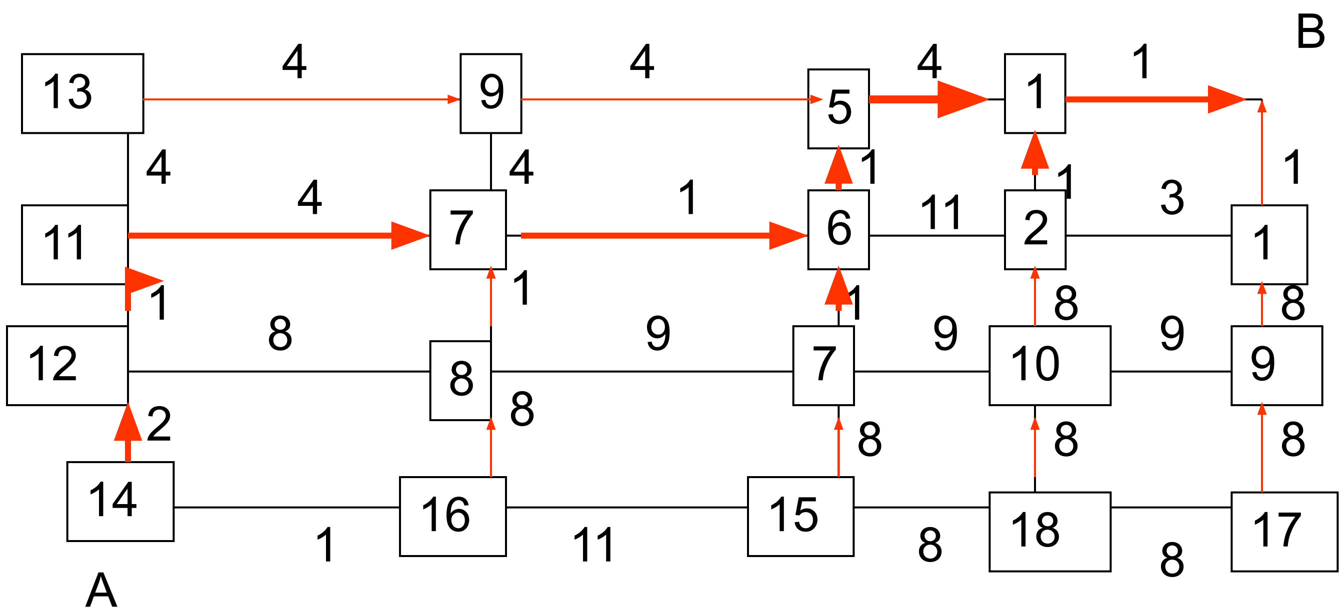
(рекуррентные соотношения, позволяющие найти предыдущее значение функции, зная последующее).

3. Задача о выборе оптимального пути

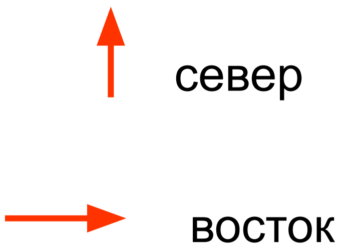
Необходимо выбрать путь из пункта А в пункт В, чтобы затраты на строительство магистрали были минимальными



Пример решения задачи динамического программирования



Оптимальное управление
Y^*=(с, с, в, в, с, в, в)



4.Задача о распределении средств между предприятиями

Двум предприятиям выделено ***а = 2000*** *единиц средств на* ***4 года***. ***Необходимо*** *распределить эти средства между предприятиями для получения* ***максимального дохода***, если ***в первый год*** средства распределяются между предприятиями в полном объеме, ***во второй*** распределяется неосвоенная за первый год часть средств (остаток) и т.д., а также известно, что

– ***доход*** от ***х*** единиц средств, вложенных на год в ***первое предприятие***, равен $f_1(x) = 6x$;

– ***доход*** от ***у*** единиц средств, вложенных на год во ***второе предприятие***, равен $f_2(y) = 4y$;

– ***остаток*** средств к концу года на ***первом предприятии*** составляет $g_1(x)=0,3x$;

– ***остаток*** средств к концу года на ***втором предприятии*** составляет $g_2(y)=0,6y$.

$f_1(x) = 6x \quad f_2(y) = 4y \quad g_1(x) = 0,3x \quad g_2(y) = 0,6y$

5.Решение задачи методом динамического программирования

Пусть в начале некоторого ***произвольного года*** мы должны распределить ***x*** единиц средств. Обозначим через ***y*** ***(0 □ y □ x)*** средства, выделяемые ***второму*** предприятию. Тогда ***первое предприятие*** получит ***x – y*** ***ед. средств.***

Обозначим $f(x,y)$ суммарный доход за этот год.

$$f(x;y) = f_1(x-y) + f_2(y) = 6(x-y) + 4y = 6x - 2y$$

Остаток средств через год обозначим $g(x,y)$

тогда $$g(x;y) = 0,3(x-y) + 0,6y = 0,3(x+y)$$

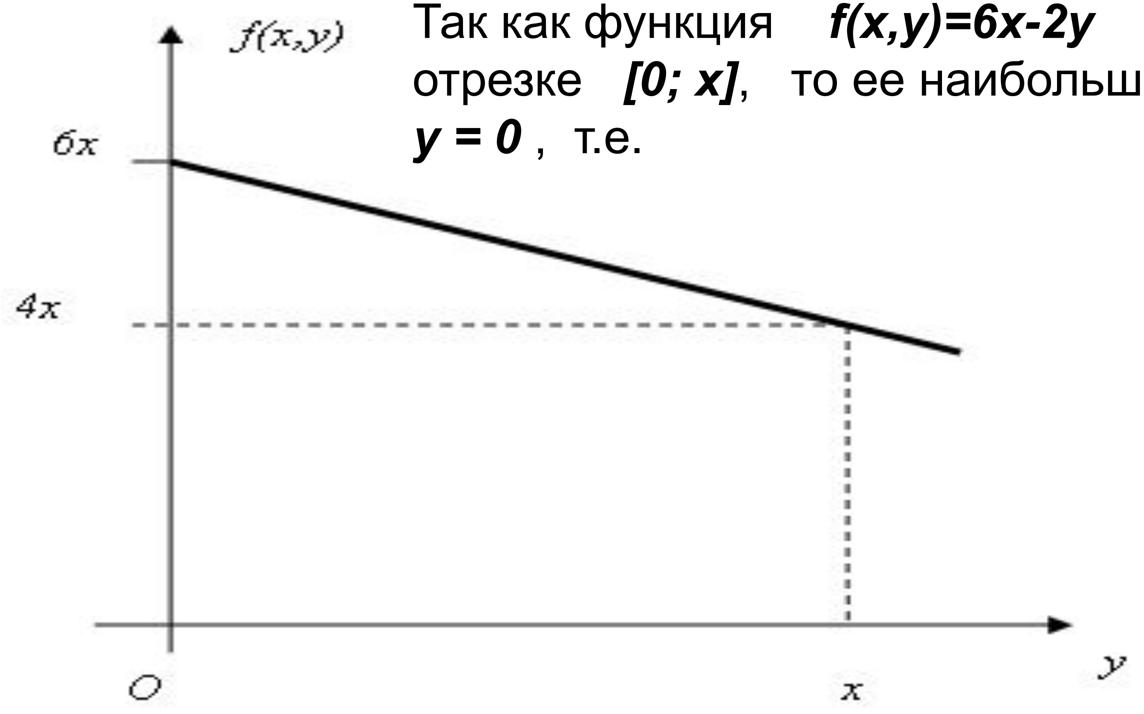
Здесь состояние системы в начале года определяется имеющимися средствами, т. е. числом ***x*** , а управление − способом распределения средств, т.е. числом ***y*** . Для состояния ***x*** при управлении ***y*** система к концу года перейдет в состояние, ***определяемое остатком средств, т.е. значением*** **g(x,y) = 0,3(x+y)**

Обозначим ***условный максимум показателя эффективности k –го шага***)

$$\varphi_k(x)$$

а условное оптимальное управление для этого состояния через ***$y_k^*(x)$***

Тогда для ***k=4*** $\varphi_4(x) = \max\limits_{0 \le y \le x} f(x;y) = \max\limits_{0 \le y \le x} (6x - 2y) = 6x$



Так как функция ***f(x,y)=6x-2y*** убывает по переменной ***y*** на отрезке ***[0; x]***, то ее наибольшее значение достигается при ***y = 0*** , т.е.

$$\varphi_4(x) = 6x, \;\; y_4^*(x) = 0$$

Здесь y_4^* − условное оптимальное управление на четвертом шаге

Для ***k=3,2,1*** справедливо рекуррентное соотношение

$$\varphi_k(x) = \max_{0 \le y \le x} \{ f(x;y) + \varphi_{k+1}(g(x;y)) \}$$

поэтому для ***k=3*** имеем

$$\varphi_3(x) = \max_{0 \le y \le x} \{6x - 2y + \varphi_4(0,3(x+y))\} =$$

$$= \max_{0 \le y \le x} \{6x - 2y + 6 \cdot 0,3(x+y)\} =$$

$$= \max_{0 \le y \le x} \{7,8x - 0,2y\}$$

Функция $z = 7,8x - 0,2y$ убывает по ***y*** на отрезке ***[0,x]*** , поэтому

$$\varphi_3(x) = 7,8x, \quad y_3^*(x) = 0$$

Для ***k=2***

$$\varphi_2(x) = \max_{0\le y\le x} \{6x - 2y + \varphi_3(0,3(x+y))\} =$$

$$= \max_{0\le y\le x} \{6x - 2y + 7,8 \cdot 0,3(x+y)\} = \max_{0\le y\le x} \{8,34x + 0,34y\}$$

Функция z=***8,34x +0,34y*** возрастает по ***y*** поэтому ее максимальное значение на отрезке ***[0, x]*** достигается при ***y=x***, т.е.

$$\varphi_2(x) = 8,34x + 0,34x = 8,68x, \quad y_2^*(x) = x$$

Для k=1

$$\varphi_1(x) = \max_{0\le y\le x} \{6x - 2y + \varphi_2(0,3(x+y))\} =$$

$$= \max_{0\le y\le x} \{6x - 2y + 8,68 \cdot 0,3(x+y)\} = \max_{0\le y\le x} \{8,604x + 0,604y\}$$

Функция z=***8,604x +0,604y*** возрастает по ***y***, поэтому

$$\varphi_1(x) = 9,208x, \quad y_1^*(x) = x$$

$$\varphi_1(a) = \varphi_1(2000) = 9,208 \cdot 2000 = 18416$$

Получили наибольший суммарный доход, который может быть получен при заданных условиях за 4 года. При этом средства следует распределять следующим образом: в первые два года все отдавать второму предприятию , а в последующие два года − первому предприятию .

$(y_1(x) = x;\ \ y_2(x) = x)$ $\qquad$ $(y_3(x) = 0;\ \ y_4(x) = 0)$