#### Лекция №13

###### 6. Вывод основного функционального уравнения динамического программирования

Пусть

*Ei* *S*  - (6.1)

- есть условный оптимальный выигрыш, получаемый на всех последующих шагах, начиная с i-го и до конца; он достигается при оптимальном управлении на этих шагах и равен максимальному выигрышу, который можно получить на всех этих шагах вместе, если в их начале система

находится в состоянии S.

Будем называть Пусть

E *S*  *условным оптимальным выигрышем*.

*ui* *S*  - (6.2)

i

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

Будем называть *ui* *S* - условным оптимальным управлением.

**Задача**: определить

*Ei* *S*  и *ui* *S*  для всех шагов (*i*=1,2,…,*m*).

Рассмотрим *i*-й шаг процесса управления.

Пусть в результате (*i-1*) предыдущих шагов система пришла в состояние *S*, и мы выбираем какое-то (вообще говоря, не оптимальное)

управление *Ui*

выигрыш:

на *i*-ом шаге. Если его применить, то получим на *i*-ом шаге

*ei*  *ei* *S,Ui* 

А система перейдёт в новое состояние *S*  :

*S*   *S*,*U* 

*i i*

(6.3)

(6.4)

Если ищется оптимальное управление на *i*-ом шаге, то согласно

принципу оптимальности Беллмана предполагается, что выигрыш

*Ei*1 (*S*) на

всех последующих шагах, начиная с *i+1* (от нового состояния *S*  ) – условный оптимальный выигрыш. А так как на *i*-ом шаге *Ui* - есть еще не оптимальное управление, то, применив его на *i*-м шаге, получим некоторый

полуоптимальный выигрыш:

~    *e* *S*,*U*

 *E*

*S* - или с учетом (6.4):

Ei *S*,*Ui i*

*i i*1

~    *e* *S*,*U*

 *E*  *S*,*U* 

(6.5)

Ei *S*,*Ui i*

*i i*1 *i i*

Теперь, в соответствии с принципом Беллмана, надо выбрать такое

*U*i  *ui* , при котором (6.5) максимальна и достигает значения:

E *S*   max*e* *S*,*U*

 *E*  *S*,*U*



(6.6)

i *Ui i*

*i i*1 *i i*

То управление

*Ui*  *ui* , при котором (6.6)=max, и есть условное

оптимальное управление на *i*-ом шаге, а величина

*Ei* (*S* )

из (6.6) – условный

оптимальный выигрыш (на всех шагах, начиная с *i*-го и до конца).

*i*

*i*

*i*

В уравнении (6.6) функции

*i*

*e* *S* ,*U*

 и  *S* ,*U*

 известны. Неизвестными

остаются функции вторую.

i

i1 *i*

E *S*  и

E *S* ,*U* . Из них первая выражается через

###### Формула (6.6) – есть основное функциональное уравнение

**динамического программирования**. Она позволяет определить

E *S* , если

известна следующая за ней по порядку функция Мы знаем, что для последнего (m-го) шага

i

i1

E *S* .

E *S*   max*e* *S*,*U* 

(6.7)

m *Um m m*

Причём, максимум в (6.7) берётся только по тем *Um* , которые приводят

систему в заданную область конечных состояний

*S* , т.е. по тем, для которых

 *S*,*U*  *S*

*m m*

* это надо иметь в виду.

То управление

*U*  *u* *S* , при котором достигается максимум

выигрыша (6.7), и есть условное оптимальное управление на *m*-м шаге.

*m m*

Зная

E *S* 

и полагая *i+*1*=m*, можно по формуле (6.6) найти

E *S*  и

*um -*1 *S* ; затем

m

m-1

*Em-*2 *S*  и

*um-*2 *S* 

и т.д., вплоть до последнего от конца

(первого) шага, для которого будут найдены функции

E1 *S*  и *u*1 *S* . Функция

E *S* 

1

* 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 *S* системы).

Таким образом, предварительная оптимизация закончена – найден условный оптимальный выигрыш и условное оптимальное управление для каждого шага.

Теперь можно перейти к нахождению *безусловного* оптимального

управления

*u*  *u*1*,u*2*,...um* .

Начнём с первого шага. Пусть *S*0 – исходное состояние. Подставим S0

в формулу для

E *S* . Получим

Emax

 *E* *S* 

1

(6.8)

и *u*  *u* *S*  .

1

0

1 1 0

Далее, зная S0 и u1, можем найти состояние S\* после первого шага

1

S\*   *S* ,*u* 

(6.9)

1 1 0 1

Зная S\* , можно найти *u*  *u* *S*\* , затем S\*   *S*\* ,*u*

 и т.д., т.е.

1 2 2 1 2 2 1 2

S  *u* *S*

  *S*\*  *u* *S*\*  ...  *S*\*

 *u* *S*\*  *S*\*

(6.10)

0 1 0

Естественно, что

1 2 1

*S*\*  *S*

*m*1 *m m*1 *m*

 ~

(6.11)

*m*  *S*

и

*u*  *u*1*,u*2*,...um* .

Итак, при решении любой задачи динамического программирования необходимо придерживаться стандартного порядка действий. Например, так:

1. Выбрать способ описания процесса, т.е. параметры, характеризующие состояние системы, фазовое пространство и способ членения операции на «шаги».
2. Записать выигрыш на *i*-ом шаге в зависимости от состояния системы S в начале этого шага и управления Ui:

*ei*  *ei* *S,Ui* .

1. Записать для *i*-го шага функцию, выражающую изменение S на

*S*  под управлением *Ui*:

S   *S*,*U* .

i *i*

1. Записать основное функциональное уравнение (6.6):

E *S*   max*e* S,*U*  *E*  *S*,*U* 

i

1. Найти

*Ui i i*

*i*1 *i i*

*E* *S*   max*e* *S*,*U* 

*m Um m m*

и соответствующее ей условное оптимальное управление на последнем шаге: *um* *S* .

*i i i i*

1. Для конкретных

*e* *S* ,*U*  и  *S* ,*U*  найти цепочку

*Em*1

*S*, *E*

*S*,..., *E* *S*.

и соответствующие им условные оптимальные управления:

*m*2

1

*um*1*S*,*um*2 *S* ,...,*u*1*S*.

1. Для начального S0 найти оптимальный выигрыш

*E*max  *E*1 (*S*0 ) и

далее безусловные оптимальные управления по цепочке:

S  *u* *S*

  *S*\*  *u* *S*\*  ...  *S*\*

 *u* *S*\*  *S*\* .

0 1 0

1 2 1

*m*1

*m m*1 *m*

1. Если начальное состояние S0 не задано, а лишь ограничено условием

S  ~

0 *S*0 ,

найти оптимальное начальное состояние достигает максимума.

0

*S*\* , при котором выигрыш

*E*1 *S* 

*E*max  max*E*1 *S* .

*S*~

*So*

и далее, по цепочке, безусловные оптимальные управления.

В заключение отметим, что в принципе процесс ДП может разворачиваться (хотя и не так естественно) и в направлении, обратном тому, которое описано: условные оптимальные управления могут отыскиваться в направлении от первого шага к последнему – а безусловные – от последнего к первому.

То есть, сначала можно планировать 1-й шаг, при условии, что он приведет систему в состояние S, а затем второй, так, чтобы выигрыш за два первые шага (первый – уже оптимизированный) был максимален, и т.д.

Результаты оптимизации управления для обеих схем будут одинаковыми

###### 7. Постановка и схема решения задачи распределения ресурсов.

Пусть

*K*0 - начальное количество средств, которые нужно распределять

в течение *m* лет между двумя предприятиями П1 и П2. Предприятия приносят за год определенный доход, зависящий от объема вложений.

Если средства Х ед. вложить в П1, то за год получим доход - *f* ( *X* ) . За

год вложенные средства частично амортизируются и к концу года от них остаётся часть

( *X* )  *X* .

Аналогично для П2: вложения *Y* ; доход за год

 (*Y* )  *Y* .

*g* (*Y* )

и остаток года

По истечении года, оставшиеся от

*K*0 средства заново распределяются

по предприятиям П1 и П2. Новых средств извне не поступает, а в производство вкладываются все оставшиеся в наличии средства; доход в производство не вкладывается, а накапливается отдельно.

Требуется найти *такой способ управления* ресурсами (т.е. какие средства, в какие годы и в какое предприятие вкладывать), при котором суммарный доход от обоих предприятий за *m* лет будет максимальным.

Решаем задачу методом ДП по известной схеме.

1. Система *S* - два предприятия с вложенными в них средствами. Она характеризуется двумя параметрами X и Y. Шагом процесса является хозяйственный год. X и Y количества средств в предприятиях П1 и П2.

В процессе X и Y меняются в зависимости от двух причин:

* + перераспределение средств в начале каждого года между предприятиями;
  + уменьшение средств за год, сказывающееся в конце каждого года.

Управлением *Ui* на *i*-м шаге будут количества средств X*i* и Y*i*,

вкладываемые в П1 и П2 на этом шаге. Таким образом, управление операцией

*U* состоит в совокупности всех шаговых управлений:

*U*  (*U*1,*U* 2 ,...,*Um* )

необходимо найти такое (оптимальное) управление

(7.1)

*u*  (*u*1, *u*2 ,...,*um* ) , (7.2)

при котором суммарной доход, приносимый за *m* лет обоими предприятиями был бы максимальным, т.е.

*E*  *E*max , где

*m*

*E*   *ei* . (7.3)

*i*1

1. Состояние *S* системы перед *i* -м шагом характеризуется одним параметром

*K* - количеством средств, сохранившихся после предыдущих

*i* 1

шагов.

Управление *Ui*

будет состоять в том, что мы для П1 выделим

*X i* средств; а

для П2 – автоматически

Доход на *i* -м шаге будет:

*Yi*  *K*  *Xi*

*ei* (*K*, *Xi* )  *f* ( *Xi* )  *g*(*K*  *Xi* ) (7.4)

1. Под влиянием этого управления на *i* -м шаге система перейдёт из состояния *K* в состояние *K* :

*S*  *K*  ( *Xi* )  (*K*  *Xi* )

1. Основное функциональное уравнение ДП имеет вид:

(7.5)

*Ei* (*K* )  max *f* ( *Xi* )  *g*(*K*  *Xi* )  *Ei*1 (( *Xi* )  (*K*  *Xi* )), (7.6)

0 *Xi* *K*

где

0  *X*  *K*

обозначает, что максимум берется по всем неотрицательным

вложениям

*Xi* , не превосходящим наличного запаса средств *K* . Условным

оптимальным управлением на *i* -м шаге

*xi* (*K* )

будет то из значений

*Xi* , при

котором

*Ei* (*K* )

достигает максимум.

1. Условный оптимальный выигрыш на последнем шаге будет:

*Em* (*K* )  max *f* ( *Xm* )  *g*(*K*  *Xm* ); (7.7)

0 *Xi* *K*

ему соответствует условное оптимальное управление этот максимум достигается.

*xm* (*K* ) , при котором

1. Зная функцию *Em* (*K* ) , находим по формуле (7.6) условные оптимальные

выигрыши на двух последних, на трёх и т.д. шагах:

*Em*1 (*K* ) 

*Em*2 (*K* ) 

max

0 *Xm*1*K*

max

*f* ( *Xm*1 )  *g*(*K*  *Xm*1 )  *Em* (( *Xm*1 )  (*K*  *Xm*1 ));

*f* ( *Xm*2 )  *g*(*K*  *Xm*2 )  *Em*1 (( *Xm*2 )  (*K*  *Xm*2 ));

.................

0 *Xm*2 *K*

(7.8)

*E*1 (*K* ) 

max *f* ( *X*1 )  *g*(*K*  *X*1 )  *E*2 (( *X*1 )  (*K*  *X*1 )).

0 *X*1*K*

Им соответствуют условные оптимальные управления

*xm*1 (*K* ), *xm*2 (*K* ),..., *x*1 (*K* )

(5.9)

1. Начальный запас средств задан, т.е. задано начальное состояние *K*0

поэтому оптимальный выигрыш будет

*E*max  *E*1 (*K*0 )

Оптимальное управление на 1-м шаге:

*x*1  *x*1 (*K*0 ) .

Состояние системы после первого шага:

*K*\*  (*X* ) (*K*  *X* )

1 1 0 1

Оптимальное управление на втором шаге:

*x*  *x* (*K* \*) ,

и т.д. по цепочке.

2 2 1

Состояние системы после *i* -шагов:

*K*\*  (*X* ) (*K*\*  *X* )

*i i i*1 *i*

Оптимальное управление на *i* -м шаге:

*x*  *x* (*K* \* ) ,

*i i i*1

и т.д. вплоть до последнего шаге, по цепочке:

*K*  *x* (*K*

)  *K* \*  *x* (*K* \*)  ...  *K* \*

 *x* (*K* \* )  *K* \*

0 1 0

1 2 1

*m*1

*m m*1 *m*

Величина *K*\* - количество средств, оставшихся (при оптимальном

*m*

управлении) после последнего шага.

Совокупность средств, вложенных по годам в П1:

*x*  (*x*1, *x*2 ,..., *xm* ) - есть оптимальное управление.

Количество средств, вложенных в П2:

*y*  ( *y* , *y* ,..., *y*

)  ((*K*

 *x* ),(*K*\*  *x* ),...,

(*K*\*  *x* ))

1 2 *m*

0 1 1 2

2 *m*1 *m*

Дадим геометрическую интерпретацию процесса распределения ресурсов.

Y Сумма средств, вкладываемых в П1 и П2 не

А может быть больше, чем начальное количество

средств

~

*K*0 , поэтому фазовое пространство –

K0 *S*0

~

*S*

это часть плоскости XOY, заключенный внутри равнобедренного треугольника AOB с катетами *K*0. Так как в начале распределения

0 В Х

сумма средств в обоих предприятиях равна

*K*0 ,

то ~ -есть гипотенуза треугольника *AB* .

*S*0

Так как на количество средств в конце периода *m* лет никаких

ограничений, кроме

0  *X*

 *Y*

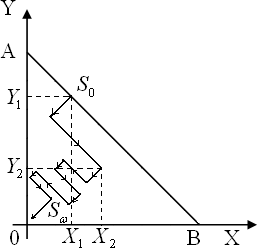
 *K*0

, не накладывается, то область

~ - есть

весь треугольник *AOB* (кроме гипотенузы).

*S*



Каждое звено траектории *S* в фазовым пространстве состоит их двух полузвеноев: на первом происходит только перераспределение средств и точка *S* перемещается параллельно *AB* , на втором – средства тратятся и точка *S* перемещается вниз и налево, ближе к началу координат. Исключение составляет первой шаг – первое полузвено отсутствует: сразу назначаются

*X*1 и *Y*1

и начинается трата средств. Сумма абсциссы и ординаты последней

точки траектории *S*

* есть количество средств

*K* , которое сохранится к

концу периода при данном управлении.

###### Контрольные вопросы

1. Выведите основное функциональное уравнение ДП.
2. Почему в ДП употребляются понятия: «условное оптимальное управление», «оптимальное управление» и, соответственно –

«условный оптимальный выигрыш» и «оптимальный выигрыш»? Чем они отличаются друг от друга?

1. Опишите стандартный порядок действий при решении любой задачи методом ДП.
2. Рассмотрите реализацию стандартный схемы ДП при решении задачи распределения ресурсов.
3. Как строится область допустимых решений (ОДР) при решении задачи распределения ресурсов между двумя предприятиями методом ДП?