#### Лекция №11

###### Динамическое программирование

###### Введение

*Динамическое программирование* (ДП) – математический метод оптимизации решений, специально приспособленный *к многошаговым (многоэтапным) операциям*.

Некоторые операции расчленяются на шаги естественно: например, при планировании хозяйственной деятельности группы предприятий естественным шагом является хозяйственный год. В других случаях разделение на шаги вводится искусственно: например, процесс вывода ракеты на орбиту можно условно разбить на этапы, каждый из которых занимает какое-то ∆*t.*

Процессы, о которых идёт речь, *управляемые*, т.е. на каждом шаге принимается какое-то *решение*, от которого зависит успех данного шага и *операции в целом*.

*Управление операцией складывается из ряда шаговых управлений.*

Рассмотрим пример *естественно-многошаговой операции*.

Пусть планируется деятельность группы предприятий П1, П2, …, П*к* на некоторый *период* Т, состоящий из m хозяйственных лет.



В начале периода на развитие системы выделяются основная средства К0, которые должны быть как-то распределены между Пi. В процессе функционирования системы средства частично расходуются (амортизируются). Кроме того, каждое Пi за год приносит доход, зависящий от вложенных средств. В начале каждого хозяйственного года имеющиеся средства могут перераспределяться между Пi.

**Ставится вопрос**: как нужно в начале каждого хозяйственного года распределять имеющиеся средства между предприятиями, чтобы суммарный доход от всей системы предприятий за весь период *T=m был максимальным*?

Это типичная задача ДП.

Рассматривается управляемый процесс – функционирование системы предприятий. *Управление* – распределение (или перераспределение) средств. *Шаг управления* – выделение каких-то средств каждому из Пi в начале хозяйственного года.

Пусть в начале i-го года предприятиям П1, П2, …, П*к* выделяются соответствующие средства:

*X* 1 *, X* 2  *,..., X* *k*  *.*

*i i i*

Совокупность

*X*  *j*  *, j*  1*,k*

* есть управление на i-м шаге:

*U*  *X* 1 *, X* 2 *,..., X* *k*  

*i*

(1.1)

*i i i i*

Управление U операцией в целом – есть совокупность всех шаговых управлений:

*U*  *U ,U ,...,U* 

1 2 *m*

(1.2)

Эффективность управления *U* оценивается тем же показателем Е, что и эффективность операции в целом. В примере Е (целевая функция) – есть

суммарный доход от всей системы предприятий за m лет:

*E*  *E**U*   *E**U ,U ,...,U* *.*

1 2 *m*

(1.3)

**Вопрос**: Как выбрать шаговые управления величина Е достигла максимального значения.

*U*1 *,U*2 *,...,Um ,*

для того, чтобы

Поставленная задача называется *задачей оптимизации управления*, а

управление, при котором *E=max*, называется *оптимальным управлением*.

###### Пошаговая оптимизация многошаговой операции – основной принцип динамического программирования

Будем обозначать оптимальное управление (в отличие от управления вообще *U*) маленькой буквой *u*.

Оптимальное управление *u* многошаговым процессом состоит из

совокупности оптимальных шаговых управлений.:

*u*  *u ,u*

1

2

*,...,um* 

(1.4)

Таким образом, стоит *задача*: определить оптимальное управление на

каждом шаге операцией *u*.

*ui ,i* 1*,*2*,...,m*

и, значит, оптимальное управление всей

Заметим, что в нашем примере (управление финансированием

предприятий) показатель *Е* представляет собой сумму доходов за все отдельные годы (шаги):

*m*

*E*  *i*

*i*1

где *i* - доход от всей *системы предприятий* за *i*-й год.

(1.5)

Показатель Е, обладающий свойством (1.5), называется *аддитивным*.

Заметим, что иногда возникают задачи, когда

*m*

*E*  *i ,*

*i*1

(1.5’)

Такой показатель Е называется *мультипликативным* (легко показать, что любой мультипликативный критерий может быть сведён к аддитивному, если прологарифмировать выражение (1.5’) и искать решение, обращающее в

максимум логарифм Е. Так как

*log E*

– возрастающая функция, то максимум

*log E*

соответствует максимуму Е).

Задача ДП в общем виде выглядит так: имеется операция с аддитивным

показателем эффективности Е, распадающаяся на *m* шагов (искусственно или

естественно). На каждом шаге применяется какое-то управление

Требуется найти оптимальное управление

*u*  *u*1,*u*2 ,...,*um* , при котором

*Ui* .

*m*

*E*  *i*

*i*1

обращается в максимум.

В методе ДП оптимальное управление *u* строится шаг за шагом, *на каждом этапе расчёта, оптимизируя только один шаг*. Такая идея

пошаговой оптимизации и *есть суть метода ДП*. Причём принцип ДП отнюдь не предполагает, что каждый шаг оптимизируется отдельно, независимо от других; что, выбирая шаговое управление, можно забыть обо всех других шагах. Напротив, шаговое управление должно выбираться *с учетом всех его последствий в будущем*.

**Например**: планируется работа группы предприятий, одни из которых выпускают предметы потребления, другие – машины для этого предприятия. Задачей является получение за m лет *максимума предметов потребления*. Пусть планируется капитальное вложение на первый год. Исходя из узких интересов этого года (шага) мы должны были бы все средства бросить на производство предметов потребления, пустить машины на полную мощность и добиться к концу года максимум предметов потребления.

Но такое решение неверно. Имея в виду будущее, необходимо выделить долю средств и на производство машин. При этом объем произведенных предметов потребления снизится, зато будут созданы условия, позволяющие увеличивать её производство в последующие годы.

Таким образом, *планируя многошаговую операцию, необходимо выбирать управление на каждом шаге с учетом его будущих последствий на ещё предстоящих шагах*.

Из этого правила есть исключение. Очевидно, *последний шаг* может планироваться без «оглядки на будущее». Его можно планировать так, чтобы он принёс наибольшую выгоду. Спланировав оптимально последний шаг, можно к нему пристраивать предпоследний, к предпоследнему – пред- предпоследний и т.д.

Поэтому процесс ДП разворачивается от конца к началу: раньше всех планируется m-й шаг. А как его спланировать, если мы не знаем, чем кончился предпоследний? Нужно сделать разные предположения о том, чем кончился предпоследний (m-1)-й шаг, и для каждого из них найти такое управление, при котором выигрыш (доход) на последнем шаге был бы максимален. Решив эту задачу, мы найдём *условное оптимальное управление*

на m – шаге, т.е. то управление, которое надо применить, если (m-1)-й шаг закончился определенным образом. Допустим, что эта процедура выполнена и для каждого исхода (*m*-1)-го шага известно условное оптимальное управление на *m*-м том шаге и соответствующий ему условный оптимальный выигрыш.

Теперь, сделав все возможные предположения об исходе (m-2)-го шага, для каждого из этих предположений найдём такое управление на (m-1)-ом шаге, чтобы выигрыш за два последних шага (из которых последний уже оптимизирован) был максимален. Далее определяется управление на (m-2)-м шаге (делая предположения об исходах (m-3)-го шага), и т.д.

Одним словом, на каждом шаге ищется такое управление, которое обеспечивает *оптимальное продолжение процесса* относительно достигнутого в данный момент состояния.

Этот принцип выбора управления называется *принципом оптимальности*.

Управление, обеспечивающее оптимальное продолжение процесса относительно заданного состояния, называется *условным оптимальным управлением на данном шаге*.

Допустим, что условное оптимальное управление на каждом шаге нам известно: мы знаем, что делать дальше, в каком бы состоянии ни был процесс к началу каждого шага. Тогда можно найти уже не «условное», а просто оптимальное управление на каждом шаге.

Действительно, если известно начальное состояние процесса S0, то надо применить условное оптимальное управление u1, выработанное для первого шага, относящееся к состоянию S0. В результате система перейдёт

*S*0  *S*1 , но для *S*1 опять известно условное оптимальное управление на

втором шаге u2 и так далее. Таким образом, найдём оптимальное управление процессом

*u*  *u*1,*u*2,...,*um* , приводящее к максимально возможному выигрышу

*E*max .

Следовательно, в процессе оптимизации управления методом ДП многошаговый процесс «проходится» дважды:

- первый раз – от конца к началу, в результате находятся условные оптимальные управления на каждом шаге и оптимальный выигрыш (тоже условный) на всех шагах, начиная с данного и до конца процесса;

- второй раз – от начала к концу, в результате чего находятся уже не условные оптимальные шаговые управления на всех шагах операции.

###### Контрольные вопросы

1. В чем состоят особенности метода ДП, как метода для поиска оптимального управления многошаговых операций?
2. Сформулируйте принцип оптимальности Беллмана.
3. Поясните содержательный смысл выражений: «условно оптимальное управление на данном шаге», «оптимальное продолжение процесса».
4. Какой показатель эффективности управления называется аддитивным? Какой – мультипликативным? Можно ли, мультипликативный показатель перевести в аддитивным?
5. Приведите пример практической задачи поиска оптимального управления, которая может быть решена методом ДП.