Динамическое программирование

Понятие о динамическом программировании

- **ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ** [ДП] — раздел математического программирования, совокупность приемов, позволяющих находить оптимальные решения, основанные на вычислении последствий каждого решения и выработке оптимальной стратегии для последующих решений.

- Процессы принятия решений, которые строятся по такому принципу, называются **многошаговыми процессами**.

Особенности математической модели задачи ДП

- задача оптимизации формулируется как конечный многошаговый процесс управления;
- целевая функция (выигрыш) является аддитивной и равна сумме целевых функций каждого шага. $Z = \sum_{i=1}^{N} z_i$
- выбор управления $\mathbf{x_k}$ на каждом шаге зависит только от состояния системы к этому шагу $\mathbf{S_{k-1}}$, и не влияет на предшествующие шаги (нет обратной связи);
- состояние системы $\mathbf{S_k}$ после каждого шага управления зависит только от предшествующего состояния системы $\mathbf{S_{k-1}}$ и этого управляющего воздействия $\mathbf{x_k}$ (отсутствие последействия) и может быть записано в виде уравнения состояния: $\mathbf{S_k = f_k(S_{k-1}, x_k)}$;
- на каждом шаге управление $\mathbf{x_k}$ зависит от конечного числа управляющих переменных, а состояние системы зависит $\mathbf{S_k}$ – от конечного числа параметров;
- оптимальное управление представляет собой вектор , определяемый последовательностью оптимальных пошаговых управлений:

  $\mathbf{X= (x^*_1, x^*_2, ..., x^*_k, ..., x^*_n)}$, число которых и определяет количество шагов задачи.

Все это вытекает из принципа оптимальности

Принцип оптимальности

- Принцип оптимальности лежит в основе метода ДП
- Впервые сформулированный в 1953 г. американским математиком Р.Э.Беллманом
- **Формулировка принципа**: каково бы ни было состояние 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 выигрыш на данном шаге.

Ричард Эрнст Беллман

- Даты жизни: 26.08.1920-19.03.1984
- Американский математик, один из ведущих специалистов в области математики и вычислительной техники.
- Получил многочисленные результаты, связанные с применением динамического программирования в разных областях математики.
- В вариационном исчислении важную роль играет функциональное уравнение Беллмана. В математических методах оптимального управления известны функция и уравнение Беллмана.
- Р. Беллман опубликовал 619 статей и 39 книг. Многие его работы переведены на русский язык.

Идея динамического программирования

- При решении задачи на каждом шаге выбирается управление, которое должно привести к оптимальному выигрышу. Если считать все шаги независимыми, тогда оптимальным управлением будет то управление, которое обеспечит максимальный выигрыш именно на данном шаге.
- Математически оптимизационная задача строится в ДП с помощью таких соотношений, которые последовательно связаны между собой:
  - например, полученный результат для одного года вводится в уравнение для следующего (или, наоборот, для предыдущего), и т. д.
- Таким образом, можно получить результаты решения задачи для любого избранного момента времени и “следовать” дальше.

Идея динамического программирования

- Общим для задач ДП является то, что переменные в модели рассматриваются не вместе, а последовательно, одна за другой.
- Строится такая вычислительная схема, когда вместо **одной задачи со многими переменными** строится **много задач с малым числом** переменных в каждой.
- Это значительно сокращает объем вычислений.
- ДП применяется для задач:
  - связанных с течением времени.
  - многошаговым процессом решения “статической” задачи.

Идея динамического программирования

- Процесс решения при этом складывается из двух этапов.
- На первом он **ведется “с конца”:** для каждого из различных предположений о том, чем кончился предпоследний шаг, находится **условное оптимальное управление** на последнем шаге, т. е. управление, которое надо применить, если предпоследний шаг закончился определенным образом.
- Такая процедура проводится до самого начала, а затем (второй раз) выполняется **от начала к концу**, в результате чего находятся уже не условные, а **действительно оптимальные шаговые управления** на всех шагах операции

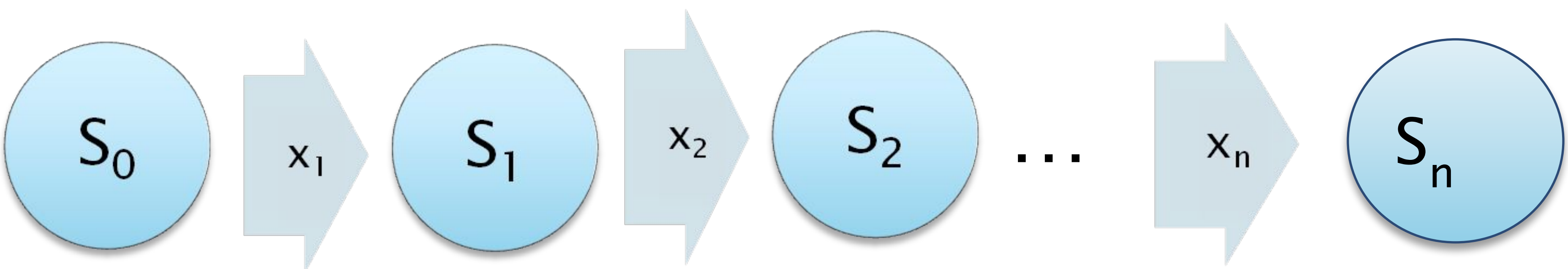
Идея динамического программирования

- В задачах, решаемых методом динамического программирования, **процесс управления разбивается на шаги (этапы).**
- При распределении на несколько лет ресурсов деятельности предприятия шагом целесообразно считать временной период;
- при распределении средств между предприятиями — номер очередного предприятия.
- В других задачах разбиение на шаги вводится искусственно. Например, непрерывный управляемый процесс можно рассматривать как дискретный, условно разбив его на временные отрезки (шаги).
- Исходя из условий каждой конкретной задачи, длину шага выбирают таким образом, чтобы на каждом шаге получить простую задачу оптимизации и обеспечить требуемую точность вычислений.

Схема решения задачи ДП

- Обозначим S_0 – начальное состояние системы, S_n - конечное.
- В результате управления система **последовательно** переводится из начального состояния S_0 в конечное S_n.
- Предположим, что управление **можно разбить на n шагов** и решение **принимается последовательно на каждом шаге**, а управление представляет собой совокупность n **пошаговых управлений**.
- На каждом шаге необходимо определить два типа переменных:
  - переменную состояния системы S_k - определяет, в каких состояниях может оказаться система на рассматриваемом *k*-м шаге;
  - переменную управления x_k - в зависимости от состояния **S** на этом шаге можно применить управления, которые удовлетворяют определенным ограничениям и называются допустимыми.
    - $X=(x_1, x_2, ..., x_k, ..., x_n)$ – общее управление, переводящее систему на состояния S_0 в состояние S_n,
      - последовательность пошаговых управлений x_k

Схема решения задачи ДП



Пример

- Владелец автомашины эксплуатирует ее в течение N лет.
- Известны затраты на содержание, ремонт и покупку нового автомобиля.
- В начале каждого года владелец может принять одно из трех решений:
  - продать машину и заменить ее новой;
  - ремонтировать ее и продолжать эксплуатацию;
  - продолжать эксплуатацию без ремонта.

- Какие нужно принять решения по годам, чтобы суммарные расходы на эксплуатацию, ремонт и приобретение новых машин были минимальны?

Пример

- Шаговое управление - выбор одного из решений на год. Припишем первому решению численное значение 1, второму 2, третьему 3.
- Показатель эффективности равен $Z = \sum_{i=1}^{N} z_i$ где – z_i расходы в i-м году.

Величину Z требуется обратить в минимум.

- Управление в целом представляет собой комбинацию чисел 1, 2, 3; например: $\vec{X} = (3,3,2,2,2,1,3,\ldots)$

Это означает, что первые два года следует эксплуатировать машину без ремонта, последующие три года ее ремонтировать, в начале шестого года продать, купить новую, затем снова эксплуатировать без ремонта и т. д.

Классические задачи ДП

- задача о выборе траектории,
- задача последовательного принятия решения,
- задача об использовании рабочей силы,
- задача управления запасами.

Оптимальная стратегия замены оборудования

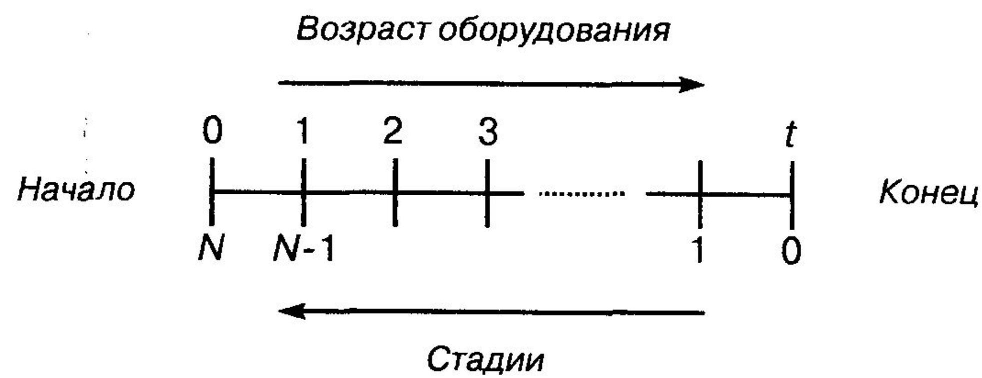
- Одной из важных экономических проблем является определение оптимальной стратегии в замене старых станков, агрегатов, машин на новые.
- Старение оборудования включает его физический и моральный износ, в результате чего растут производственные затраты по выпуску продукции на старом оборудовании, увеличиваются затраты на его ремонт и обслуживание, снижаются производительность и ликвидная стоимость.
- Наступает время, когда старое оборудование выгоднее продать, заменить новым, чем эксплуатировать ценой больших затрат; причем его можно заменить новым оборудованием того же вида или новым, более совершенным.
- Оптимальная стратегия замены оборудования состоит в определении оптимальных сроков замены. Критерием оптимальности при этом может служить прибыль от эксплуатации оборудования, которую следует оптимизировать, или суммарные затраты на эксплуатацию в течение рассматриваемого промежутка времени, подлежащие минимизации.

Оптимальная стратегия замены оборудования

- Введем обозначения:
- ***r(t)*** — стоимость продукции, производимой за один год на единице оборудования возраста *t* лет;
- ***u(t)*** — ежегодные затраты на обслуживание оборудования возраста *t* лет;
- ***s(t)*** — остаточная стоимость оборудования возраста *t* лет;
- ***P*** — покупная цена оборудования.
- Рассмотрим период ***N*** лет, в пределах которого требуется определить оптимальный цикл замены оборудования.
- Обозначим через $f_N(t)$ максимальный доход, получаемый от оборудования возраста ***t*** лет за оставшиеся *N* лет цикла использования оборудования при условии оптимальной стратегии.

Оптимальная стратегия замены оборудования

- Возраст оборудования отсчитывается в направлении течения процесса. Так, $t = 0$ соответствует случаю использования нового оборудования. Временные же стадии процесса нумеруются в обратном направлении по отношению к ходу процесса. Так, $N = 1$ относится к одной временной стадии, остающейся до завершения процесса, а $N = N$ — к началу процесса.



- На каждом этапе N-стадийного процесса должно быть принято решение о сохранении или замене оборудования. Выбранный вариант должен обеспечивать получение максимальной прибыли.

Оптимальная стратегия замены оборудования

- Функциональные уравнения, основанные на принципе оптимальности, имеют вид

$$f_N(t) = \max \begin{cases} r(t) - u(t) + f_{N-1}(t+1) & \longrightarrow \text{Сохранение;} \\ s(t) - P + r(0) - f_{N-1}(1) & \longrightarrow \text{Замена,} \end{cases}$$

$$f_1(t) = \max \begin{cases} r(t) - u(t) & \longrightarrow \text{Сохранение;} \\ s(t) - P + r(0) - u(0) & \longrightarrow \text{Замена.} \end{cases}$$

- Уравнение (1) описывает *N*-стадийный процесс, а (2) — одностадийный. Оба уравнения состоят из двух частей: верхняя строка определяет доход, получаемый при сохранении оборудования; нижняя — доход, получаемый при замене оборудования и продолжении процесса работы на новом оборудовании

Пример

- Определить оптимальный цикл замены оборудования при следующих исходных данных: $P = 10$, $S(t) = 0$, $f(t) = r(t)-u(t)$, представленных в таблице

N	0	1	2	3	4	5	6	7	8	9	10	11	12
$f(t)$	10	9	8	7	6	5	4	3	2	1	0	0	0

Пример

$$f_N(t) = \max \begin{cases} f(t) + f_{N-1}(t+1), \\ -p + f(0) + f_{N-1}(1), \end{cases}$$

$$f_1(t) = \max \begin{cases} f(t), \\ -p + f(0). \end{cases}$$

Для $N = 1$

$$f_1(0) = \max \begin{cases} f(0) \\ -p + f(0) \end{cases} = \max \begin{cases} 10 \\ -10 + 10 \end{cases} = 10,$$

$$f_1(1) = \max \begin{cases} f(1) \\ -p + f(0) \end{cases} = \max \begin{cases} 9 \\ -10 + 10 \end{cases} = 9,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$f_1(12) = \max \begin{cases} f(12) \\ -p + f(0) \end{cases} = \max \begin{cases} 0 \\ -10 + 10 \end{cases} = 0.$$

Пример

- Для N=2

$$f_2(0) = \max \begin{cases} f(0) + f_1(1) \\ -p + f(0) + f_1(1) \end{cases} = \max \begin{cases} 10 + 9 \\ -10 + 10 + 9 \end{cases} = 19,$$

$$f_2(1) = \max \begin{cases} f(1) + f_1(2) \\ -p + f(0) + f_1(1) \end{cases} = \max \begin{cases} 9 + 8 \\ -10 + 10 + 9 \end{cases} = 17,$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

Пример

$F_N(t)$	0	1	2	3	4	5	6	7	8	9	10	11	12
	N	$N-1$	11	10	9	8	7	6	5	4	3	2	1
$f_1(t)$	10	9	8	7	6	5	4	3	2	1	0	0	0
$f_2(t)$	19	17	15	13	11	9	9*						
$f_3(t)$	27	24	21	18	17	17*							
$f_4(t)$	34	30	26	24	24*								
$f_5(t)$	40	35	32	31	30	30*							
$f_6(t)$	45	41	39	37	36	35	35*						
$f_7(t)$	51	48	45	43	41	41*							
$f_8(t)$	58	54	51	48	48*								
$f_9(t)$	64	60	56	55	54	54*							
$f_{10}(t)$	70	65	63	61	60	60*							
$f_{11}(t)$	75	72	69	67	66	65	65*						
$f_{12}(t)$	82	78	75	73	72	72*							

Оптимальное распределение ресурсов

- Пусть имеется некоторое количество ресурсов *x*, которое необходимо распределить между *n* различными предприятиями, объектами, работами и т.д. так, чтобы получить максимальную суммарную эффективность от выбранного способа распределения.
- Введем обозначения:
- x_i — количество ресурсов, выделенных *i*-му предприятию;
- $g_i(x_i)$ — функция полезности, в данном случае это величина дохода от использования ресурса x_i, полученного *i*-м предприятием;
- $f_k(x)$ — наибольший доход, который можно получить при использовании ресурсов *x* от первых *k* различных предприятий.

Оптимальное распределение ресурсов

- Сформулированную задачу можно записать в математической форме:

$$f_n(x) = \max \sum_{i=1}^{n} g_i(x_i)$$

при ограничениях:

$$\sum_{i=1}^{n} x_i = x,$$

$$x_i \geqslant 0, \quad i = \overline{1, n}.$$

- Для решения задачи необходимо получить рекуррентное соотношение, связывающее $f_k(x)$ и $f_{k-1}(x)$.
- Обозначим через x_k количество ресурса, используемого k-м способом ($0 \leq x_k \leq x$), тогда для (k-1) способов остается величина ресурсов, равная (x-x_k). Наибольший доход, который получается при использовании ресурса (x-x_k) от первых (k-1) способов, составит $f_{k-1}(x-x_k)$.
- Для максимизации суммарного дохода от k-го и первых (k-1) способов необходимо выбрать x_k таким образом, чтобы выполнялись соотношения

$$f_1(x) = g_1(x),$$

$$f_k(x) = \max\{g_k(x_k) + f_{k-1}(x - x_k)\}, \quad k = \overline{2, n}.$$

Оптимальное распределение ресурсов

- Совет директоров фирмы рассматривает предложения по наращиванию производственных мощностей для увеличения выпуска однородной продукции на четырех предприятиях, принадлежащих фирме.
- Для расширения производства совет директоров выделяет средства в объеме 120 млн р. Траншами по 20 млн р. Прирост выпуска продукции на предприятиях зависит от выделенной суммы, его значения представлены предприятиями и содержатся в таблице.

Выделяемые средства, млн р.	Прирост выпуска продукции, млн р.			
	Предприятие № 1	Предприятие № 2	Предприятие № 3	Предприятие № 4
20	8	10	12	11
40	16	20	21	23
60	25	28	27	30
80	36	40	38	37
100	44	48	50	51
120	62	62	63	63

- Найти распределение средств между предприятиями, обеспечивающее максимальный прирост выпуска продукции, причем на одно предприятие можно осуществить не более одной инвестиции.

Пример

- Рекуррентные соотношения будут иметь вид:

для предприятия № 1 $f_1(x) = g_1(x_1)$,

для всех остальных предприятий

$$f_k(x) = \max\{g_k(x_k) + f_{k-1}(x - x_k)\}, \quad k = \overline{2, n}.$$

- Решение будем проводить согласно рекуррентным соотношениям в четыре этапа.
- ***1-й этап.*** Инвестиции производим только первому предприятию. Тогда

$$f_1(20) = 8, \qquad f_1(40) = 16, \qquad f_1(60) = 25,$$

$$f_1(80) = 36, \qquad f_1(100) = 44, \qquad f_1(120) = 62.$$

Пример

- **2-й этап.** Инвестиции выделяем первому и второму предприятиям. Рекуррентное соотношение для 2-го этапа имеет вид

$$f_2(x) = \max\{g_2(x_2) + f_1(x - x_2)\}.$$

Тогда

- при x = 20 $f_2(20)$ = max (8+0, 0+10) = max (8, 10) = 10,
- при x = 40 $f_2(40)$ = max (16, 8+10, 20) = max (16, 18, 20) =20,
- при x = 60 $f_2(60)$ = max (25, 16+10, 8+20, 28) = max (25, 26, 28, 28) =28,
- при x = 80 $f_2(80)$ = max (36, 25+10,16+20, 8+28, 40) = max (36, 35, 36, 36, 40) = 40,
- при x = 100 $f_2(100)$ = max (44, 36+10, 25+20,16+28, 8+40, 48) =
- max (44, 46, 45, 44, 48, 48) = 48,
- при x = 120 $f_2(120)$ = max (62, 44+10, 36+20, 25+28,16+40, 8+48, 62) = max (62, 54, 56, 53, 56, 56, 62) = 62.

Пример

- **3-й этап.** Финансируем 2-й этап и третье предприятие. Расчеты проводим по формуле

Тогда

- при $x = 20$ $f_3(20) = \max(10, 12) = 12$,
- при $x = 40$ $f_3(40) = \max(20, 10+12, 21) = \max(20, 22, 21) = 22$,
- при $x = 60$ $f_3(60) = \max(28, 20+12, 10+21, 27) = \max(28, 32, 31, 27) = 32$,
- при $x = 80$ $f_3(80) = \max(40, 28+12, 20+21, 10+27, 38) = \max(40, 40, 41, 37, 38) = 41$,
- при $x = 100$ $f_3(100) = \max(48, 40+12, 28+21, 20+27, 10+38, 50) =$
- $\max(48, 52, 49, 47, 48, 50) = 52$,
- при $x = 120$ $f_3(120) = \max(62, 48+12, 40+21, 28+27, 20+38, 10+50, 63) = \max(62, 60, 61, 55, 58, 60, 63) = 63$.

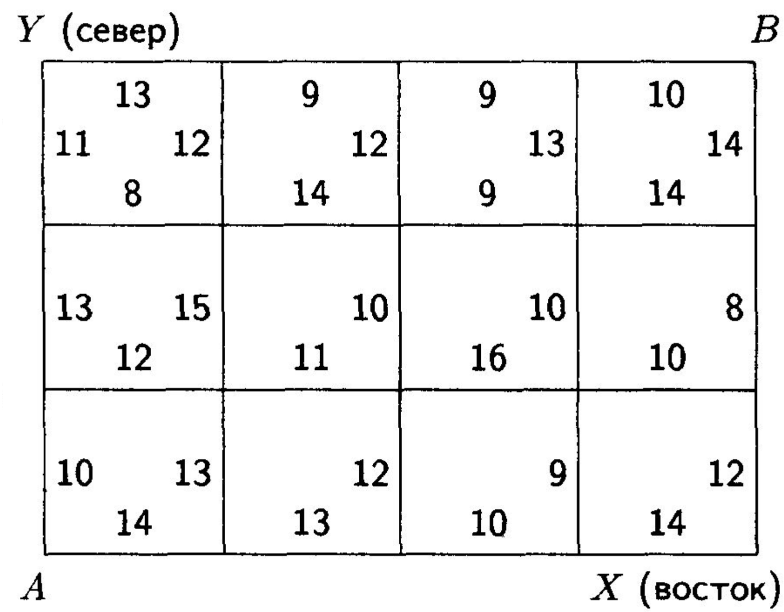
Пример

- **4-й этап.** Инвестиции в объеме 120 млн р. распределяем между 3-м этапом и четвертым предприятием.
- При x = 120 $f_4(120)$ = max (63, 52+11, 41+23, 32+30, 22+37, 12+51, 63) = max (63, 63, 64, 62, 59, 63, 63) = 64.

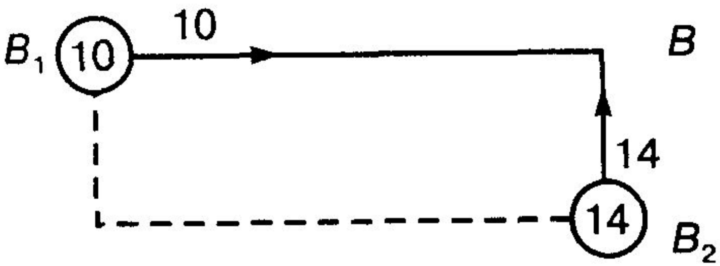
- Получены условия управления от 1-го до 4-го этапа. Вернемся от 4-го к 1-му этапу. Максимальный прирост выпуска продукции в 64 млн р. получен на 4-м этапе как 41+23, т.е. 23 млн р. соответствуют выделению 40 млн р. четвертому предприятию. Согласно 3-му этапу 41 млн р. получен как 20 + 21, т.е. 21 млн р. соответствует выделеник 40 млн р. третьему предприятию. Согласно 2-этапу 20 млн р. получено при выделении 40 млн р. второму предприятию.
- Таким образом, инвестиции в объеме 120 млн р. целесообразно выделить второму, третьему и четвертому предприятиям по 40 млн р. каждому, при этом прирост продукции будет максимальным и составит 64 млн р.

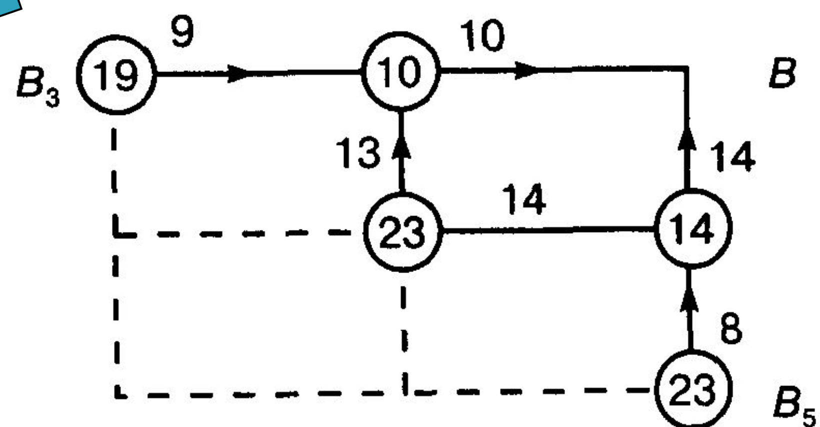
Оптимизация затрат при строительстве транспортных артерий

- Требуется проложить путь (трубопровод, шоссе) между двумя пунктами *A* и *B* таким образом, чтобы суммарные затраты на его сооружение были минимальные. Разделим расстояние между пунктами *A* и *B* на шаги (отрезки). На каждом шаге можем двигаться либо строго на восток, либо строго на север. Тогда путь от *A* в *B* представляет ступенчатую ломаную линию. Затраты на сооружение каждого из отрезков известны



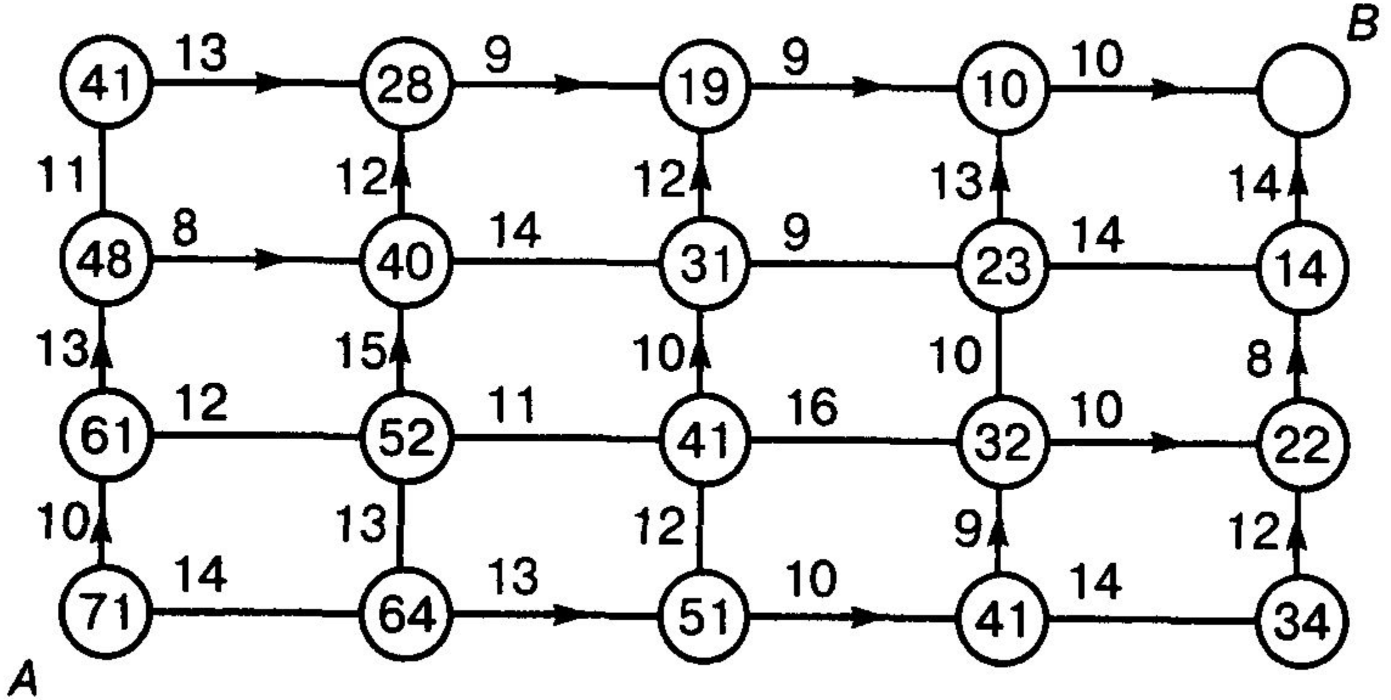
Оптимизация затрат при строительстве транспортных артерий





...

Оптимизация затрат при строительстве транспортных артерий



Преимущества и недостатки ДП

- К числу положительных качеств можно отнести:
  - МДП дает возможность решать задачи, которые раньше не исследовались из-за отсутствия соответствующего математического аппарата.
  - МДП в ряде случаев сокращает объем при поиске оптимальных решений.
- К недостаткам относятся:
  - Отсутствие универсального алгоритма, который был бы пригоден для решения всех задач (мы имеем только схему).
  - Трудности при решении задач большой размерности.