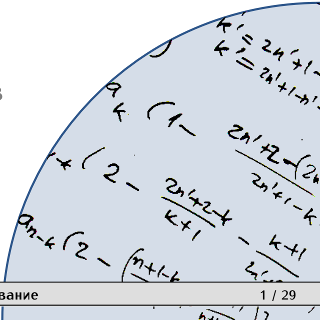
ФГБОУ ВО Вятский государственный университет
Факультет экономики и финансов

Введение в экономико-математическое моделирование

Лекция 7. Динамическое программирование

канд. физ.-матем. наук, доцент Д. В. Чупраков

usr10381@vyatsu.ru



Структура лекции

Динамическое программирование

Динамическое программирование — это метод нахождения оптимальных решений в задачах с многоэтапной структурой.

Многие экономические процессы расчленяются на шаги естественным образом. Это все процессы планирования и управления, развивающиеся во времени.

- из года в год меняется возраст машин и оборудования;
- трудозатраты меняются от работы к работе в рамках сетевой модели;
- руководство предприятием требует принятия решений в зависимости от свершившихся событий.

Очевидно, в таких задачах необходимо принимать оптимальные решения не только на текущий момент, но и на весь рассматриваемый период в целом с учетом возможных изменений параметров.

ДП связано с именем Ричарда Беллмана, который сформулировал принцип, позволяющий существенно сократить перебор решений в многоэтапных нелинейных задачах.

Управляемая система без памяти

- Конечность числа шагов:
  В результате управления система S переводится из начального состояния S_0 в состояние S_n.
- На каждом шаге $k = \{1, 2, \ldots, n\}$ принимается допустимое управляющее решение x_k.
- Система без памяти: $S_k = \varphi(S_{k-1}, x_k)$
  Состояние системы S_k в конце k-го шага определяется:
  - предшествующим состоянием S_{k-1}
  - управлением x_k
  - не зависит от других состояний и управлений

Управляемая система: $S_0 \xrightarrow{x_1} S_1 \xrightarrow{x_2} S_2 \xrightarrow{x_3} \ldots \xrightarrow{x_{n-1}} S_{n-1} \xrightarrow{x_n} S_n.$

Управление: $X = (x_1, x_2, \ldots, x_{n-1}, x_n)$

Постановка задачи ДП

Показатель эффективности управления — целевая функция

$$F = F(S_0, X)$$

Предположим, F имеет свойство аддитивности:

$$F = \sum_{k=1}^{n} f_k(S_{k-1}, x_k) = f_1(S_0, x_1) + f_2(S_1, x_2) + \ldots + f_n(S_{n-1}, x_n)$$

f_k — показатель эффективности управления на k–м шаге.

Общая формулировка задачи динамического программирования

Определить такое допустимое управление $X = (x_1, x_2, \ldots, x_{n-1}, x_n)$, переводящее систему S из состояния S_0 в состояние S_n, при котором целевая функция $F = F(S_0, X)$ принимает наибольшее значение.

Особенности модели ДП

1. Задача оптимизации интерпретируется как n-шаговый процесс управления.
2. Целевая функция (ЦФ) — сумма ЦФ на каждом шаге.
3. Выбор управления на k-ом шаге зависит только от состояния системы к этому шагу и не влияет на предшествующие шаги (нет обратной связи).
4. Состояние системы после k-го шага управления зависит от предшествующего состояния и управления.
5. На каждом шаге управление x_k зависит от конечного числа управляющих переменных, а состояние S_k — от конечного числа параметров.

Вопрос:

Что такое оптимальность управления на k–м шаге?

Условный оптимальный выигрыш $W_k(S)$ — величина выигрыша от k–го шага и до конца, если k–ый шаг начинается с некоторого состояния S.

$$F_k(S_{k-1}) = f_k(S_{k-1}, x_k) + f_{k+1}(S_k, x_{k+1}) + \ldots + f_n(S_{n-1}, x_n)$$

Принцип оптимальности Беллмана

Принимая решение k–ом шаге, нужно выбрать управление x_k так, чтобы условный оптимальный выигрыш $F_k(S_{k-1})$ был максимальным.

Ограничение: Управление на данном шаге не должно оказывать влияние на предшествующие шаги.

Последний шаг управления

- S_{n-1} — состояние системы к началу n-го шага;
- S_n — конечное состояние;
- $X_n = x_n$ — допустимое управляющее воздействие;
- $f_n(S_{n-1}, x_n)$ — целевая функция n-го шага.

По принципу оптимальности $F_n(S_{n-1}, X_n) = f_n(S_{n-1}, x_n) \to \max$.

- $F_n^*(S_{n-1})$ — максимум целевой функции n-го шага если система была в состоянии S_{n-1}.

$$F_n^*(S_{n-1}) = \max_{x_n} \left\{ f_n(S_{n-1}, x_n) \right\}$$

- $x_n^*(S_{n-1})$ — условно оптимальное управление на n-ом шаге — решение x_n, при котором достигается $F_n^*(S_{n-1})$.

Решив задачу нахождения максимума функции одной переменной x_n найдем $F_n^*(S_{n-1})$ и $x_n^*(S_{n-1})$.

В силу аддитивности

$$F_{n-1}(S_{n-2}, X_{n-1}) = f_{n-1}(S_{n-2}, x_{n-1}) + F^*(S_{n-1})$$

для $X_{n-1} = (x_{n-1}, x_n^*)$, где x_n^* — оптимальное управление n–го шага.

По принципу оптимальности $F_{n-1}(S_{n-2}) \to \max$.

$$F^*_{n-1}(S_{n-2}) = \max_{x_{n-1}}\{f_{n-1}(s_{n-2}, x_{n-1}) + F^*_n(S_{n-1})\} =$$

$$= \max_{x_{n-1}} \left\{ f_{n-1}(S_{n-2}, x_{n-1}) + F^*_n\big(\varphi(S_{n-2}, x_{n-1})\big) \right\}$$

т. к. $S_{n-1} = \varphi(S_{n-2}, x_{n-1})$.

Решив задачу нахождения максимума функции одной переменной x_{n-1} найдем

$$F^*_n(S_{n-1}) \quad \text{и} \quad X^*_{n-1} = (x^*_{n-1}(S_{n-2}), x^*_n(S_{n-1})).$$

Рекуррентное соотношение Беллмана

$F_k^*(S_{k-1})$ — условный максимум целевой функции, полученной на $n - k + 1$ шагах, начиная с k-го до конца

Основное рекуррентное соотношение Беллмана

$$F_k^*(S_{k-1}) = \max_{x_k} \left\{ f_k(s_{k-1}, x_k) + F_{k+1}^*(S_k) \right\}$$

$$F_{k+1}(S_k) = \max_{x_i} \sum_{i=k+1}^{n} f_i(S_{i-1}, x_i)$$

$$S_k = \varphi(S_{k-1}, x_k)$$

Схема применения метода ДП

1. Рассматриваем последний шаг и находим условный максимум целевой функции для каждого S_{n-1}

$$F_n^*(S_{n-1}) = \max_{x_n} \left\{ f_n(s_{n-1}, x_n) \right\}$$

2. Двигаясь с конца, для каждого k находим условные максимумы целевой функции за $n - k + 1$ шагов, начиная с k-го до конца по всем возможным управлениям x_k:

$$F_k^*(S_{k-1}) = \max_{x_k} \left\{ f_k(S_{k-1}, x_k) + F_{k+1}^*(S_k, x_{k+1}) \right\},$$

   Получаем последовательности условных оптимумов и условно оптимальных решений соответственно:

$$\begin{array}{ccccc} F_n^*(S_{n-1}) & F_{n-1}^*(S_{n-2}) & \cdots & F_2^*(S_1) & F_1^*(S_0) \\ x_n^*(S_{n-1}) & x_{n-1}^*(S_{n-2}) & \cdots & x_2^*(S_1) & x_1^*(S_0) \end{array}$$

3. Искомый оптимум целевой функции — $F_1^*(S_0)$; оптимальное решение — $X^* = (x_1^*, x_2^*, \dots, x_n^*)$.

Некоторые задачи, решаемые методом динамического программирования

Поиск оптимального маршрута

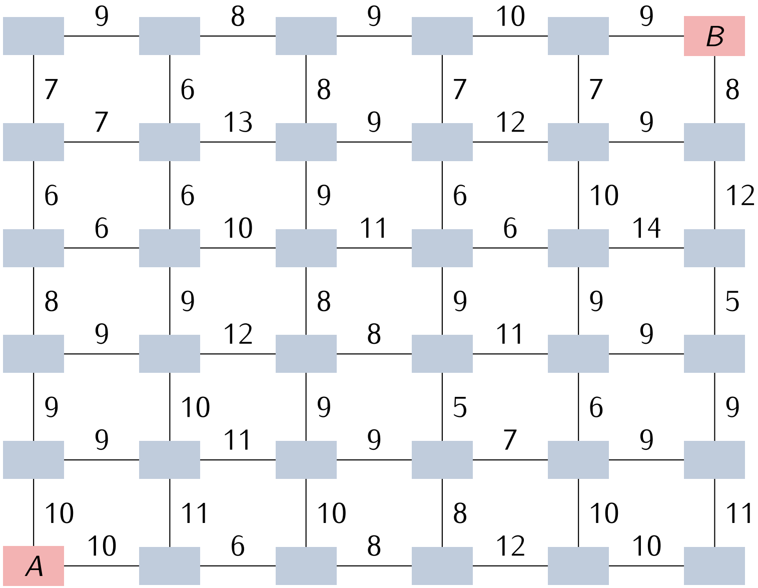
Прокладывается участок дороги из пункта A в пункт B по пересеченной местности. Требуется провести дорогу, чтобы суммарные затраты на сооружение участка были минимальные.

Формализация:

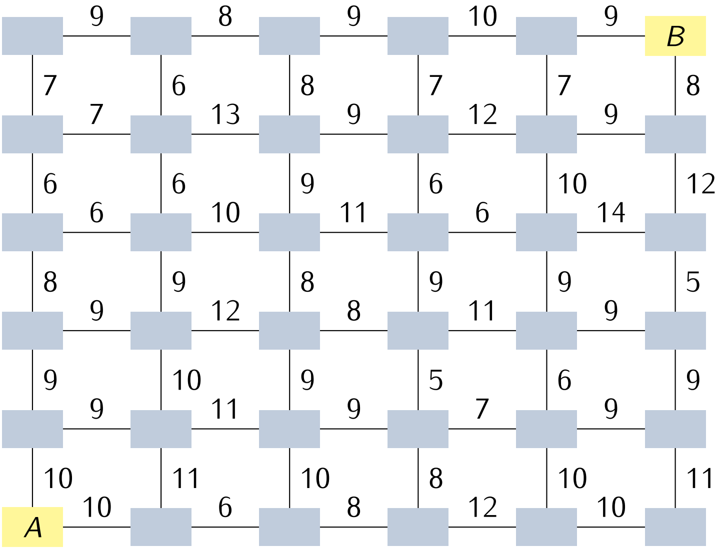
- Участок местности разбивается на сеть узлов.
- Узлы соединяются дугами, веса w_i которых обозначают затраты на строительство данного куска дороги.
- Требуется в полученном графе найти путь наименьшей стоимости.

$$W = \sum w_i \to \min$$

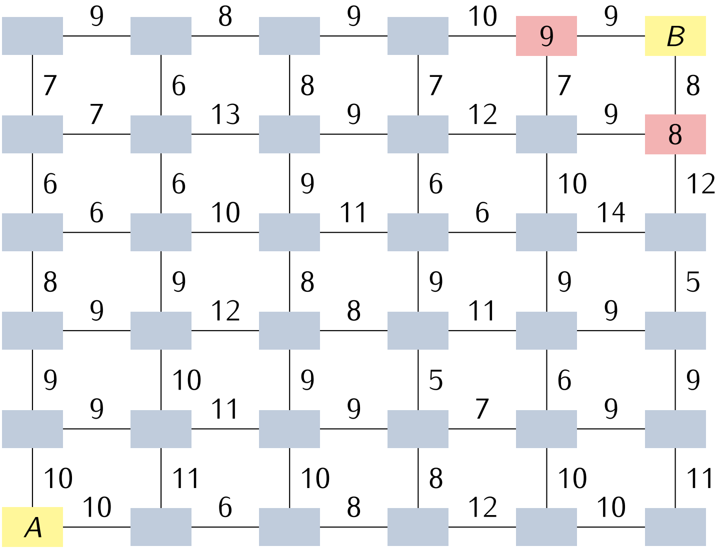
Поиск оптимального маршрута. Граф



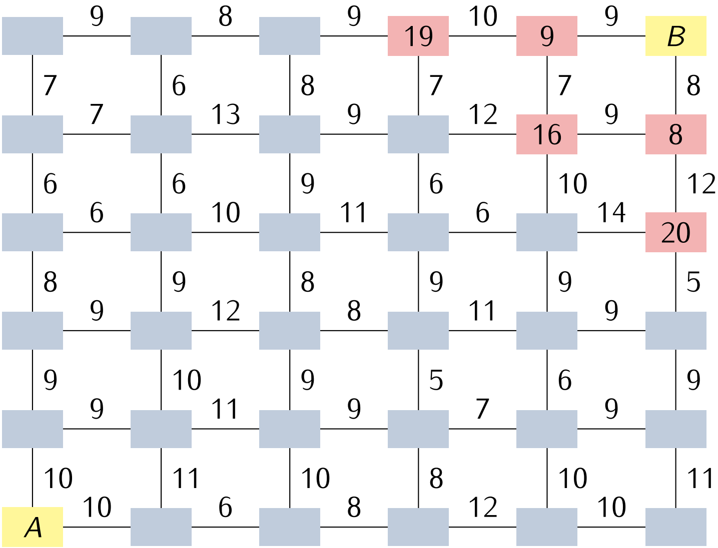
Поиск оптимального маршрута. Решение



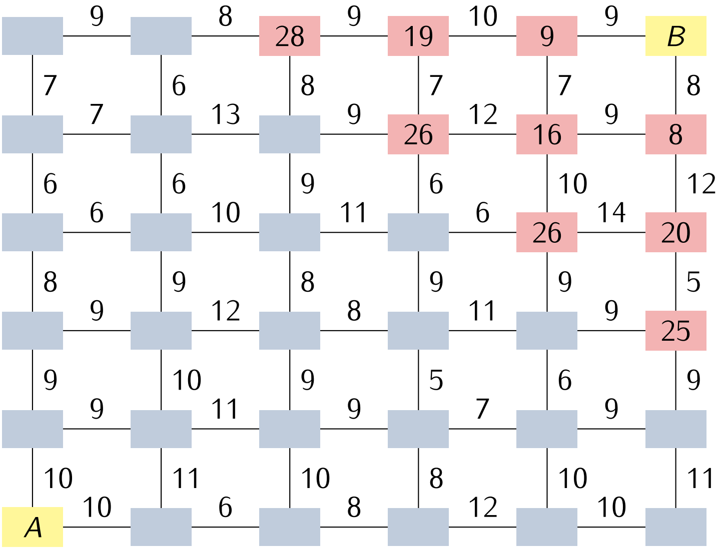
Поиск оптимального маршрута. Решение



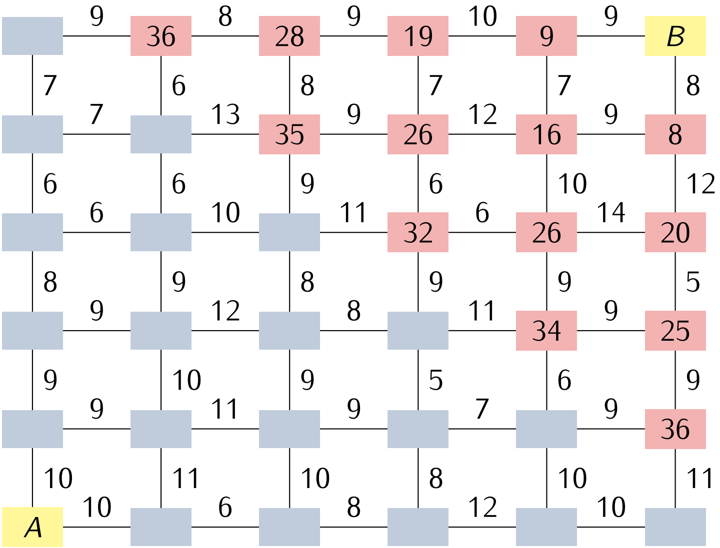
Поиск оптимального маршрута. Решение



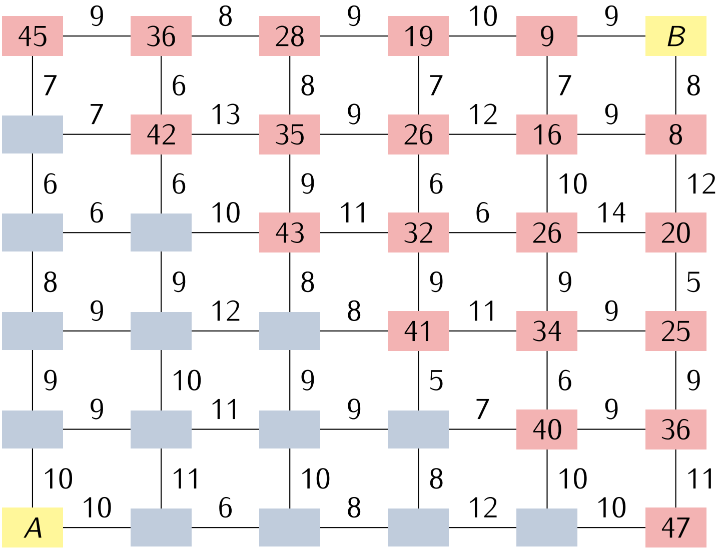
Поиск оптимального маршрута. Решение



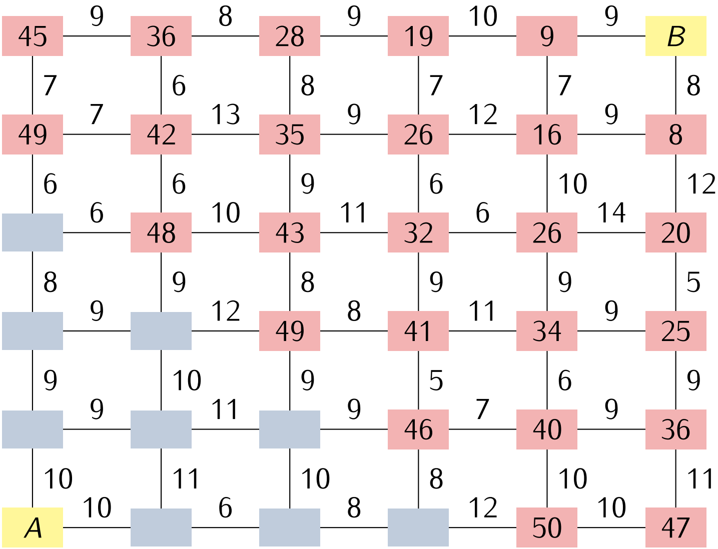
Поиск оптимального маршрута. Решение



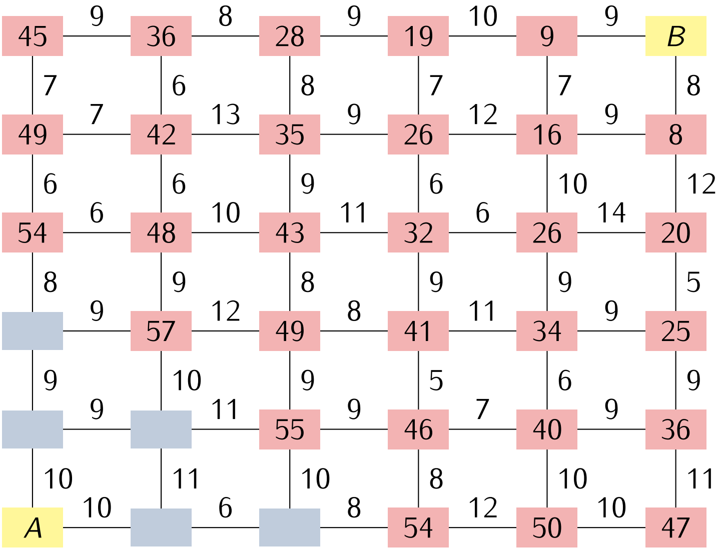
Поиск оптимального маршрута. Решение



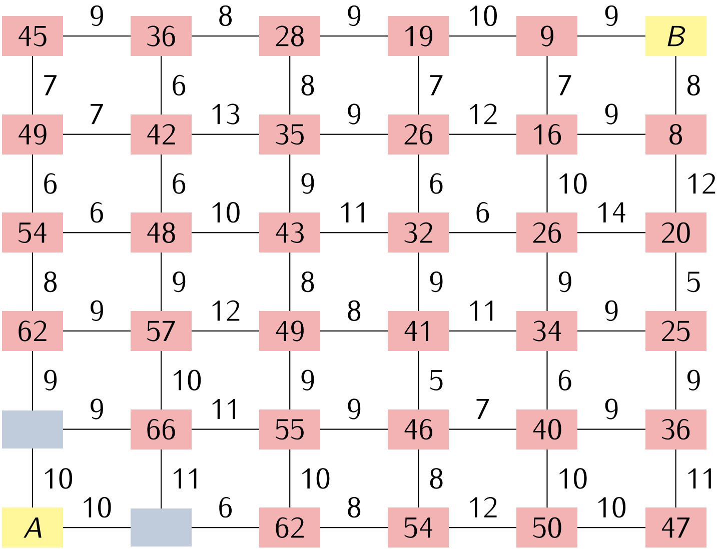
Поиск оптимального маршрута. Решение



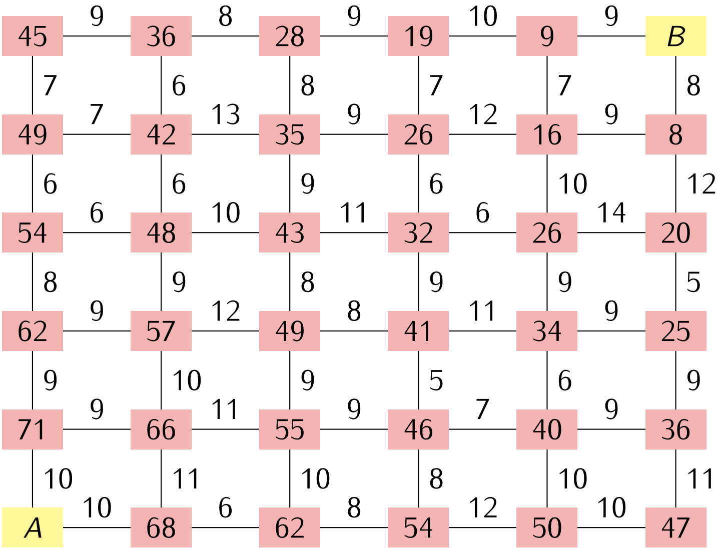
Поиск оптимального маршрута. Решение



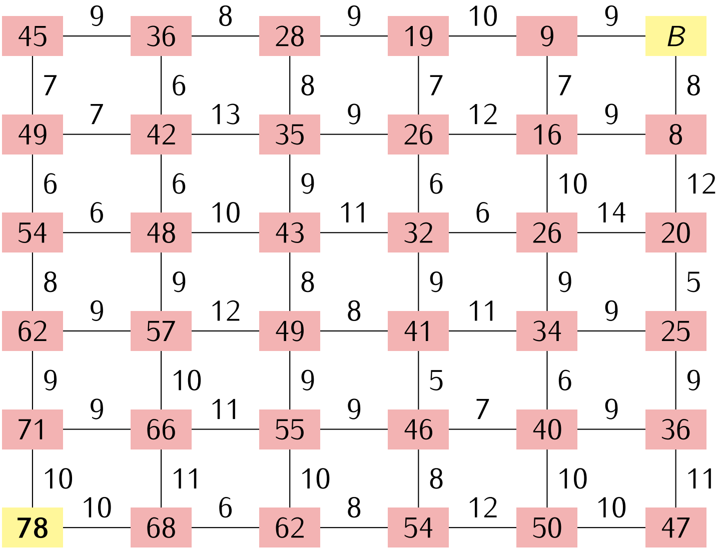
Поиск оптимального маршрута. Решение

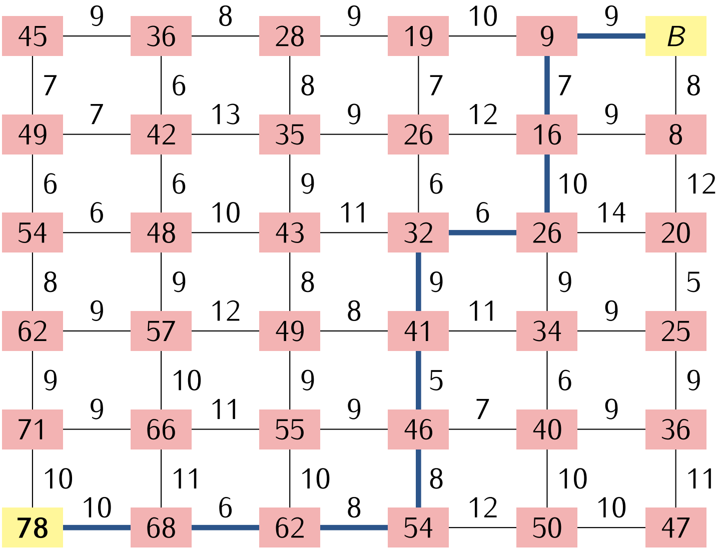


Поиск оптимального маршрута. Решение



Поиск оптимального маршрута. Решение




45 9 36 8 28 9 19 10 9 9 B
7 6 8 7 7 8
49 7 42 13 35 9 26 12 16 9 8
6 6 9 6 10 12
54 6 48 10 43 11 32 6 26 14 20
8 9 8 9 9 5
62 9 57 12 49 8 41 11 34 9 25
9 10 9 5 6 9
71 9 66 11 55 9 46 7 40 9 36
10 11 10 8 10 11
78 10 68 6 62 8 54 12 50 10 47

Ответ и структура решения

- Обратным ходом найдены общие затраты на строительство дороги — 78 условных единиц
- Прямым проходом „по карте“ находим оптимальный план: $X =$(4 вправо, 3 вверх, 1 вправо, 2 вверх, 1 вправо).

Ресурс — величина, которую система использует для производства полезного продукта.

Например:

- деньги
- время
- ГСМ
- Объем склада

Ресурс ограничен!

> Как распределить ресурс между отдельными элементами системы, чтобы суммарный эффект был максимальным?

Задача экономического планирования

- Пусть есть начальный капитал K.
- Его можно потратить на предприятия $P_1, P_2, \ldots, P_n$
- Каждое предприятие работает в течении m лет.
- X_{it} — количество средств вкладываемых в t–ом году, в i–ое предприятие.
- $f_i(X_{it})$ — доход предприятия i за год t, зависящий от вложенных средств
- Средства тратятся, принося доход, а новых средств не поступает и полученный доход не вкладывается.

Требуется так распределить капитал между предприятиями, чтобы суммарный доход был максимален:

$$F = \sum f_i(X_{ij}) \to \max, \qquad \sum X_{ij} \leqslant K$$

Распределение ресурсов на 2 предприятия

- K — начальный капитал.
- Его можно потратить на предприятия P_1 и P_2
- Каждое предприятие работает в течении m лет.
- X_t — вложения в t-ом году, в предприятие P_1.
- Y_t — вложения в t-ом году, в предприятие P_2.
- $f(X_t)$ — доход предприятия P_1 за год t,
- $g(Y_t)$ — доход предприятия P_2 за год t,
- Целевая функция — суммарный доход

$$W = \sum_{t=1}^{m} \big(f(X_t) + g(Y_t)\big) \to \max,$$

Планирования поддержки 2 предприятий

- Состоянием системы является количество средств k_t в конце t–го года.
- Управление Y_t может быть записано как $Y_t = k_{t-1} - X_t$.
- Функции возврата. В конце t–го года
  - в первой отрасли остаются средства $\varphi(X_i)$;
  - во второй — $\psi(Y_t) = \psi(k_{t-1} - X_t)$.

Составим рекурсивное соотношение Беллмана

$$W_t^*(k_{t-1}) = \max_{X_t} \left\{ f(X_t) + g(k_{t-1} - X_t) + W_{t+1}^*\Big(\varphi(X_t) + \psi(k_{t-1} - X_t)\Big) \right\}$$

Движемся с конца к началу

На последнем шаге $t = m$:

$$W_m^*(k_{m-1}) = \max_{X_m} \left\{ f(X_t) + g(k_{m-1} - X_m) \right\}$$

На предпоследнем шаге $t = m - 1$:

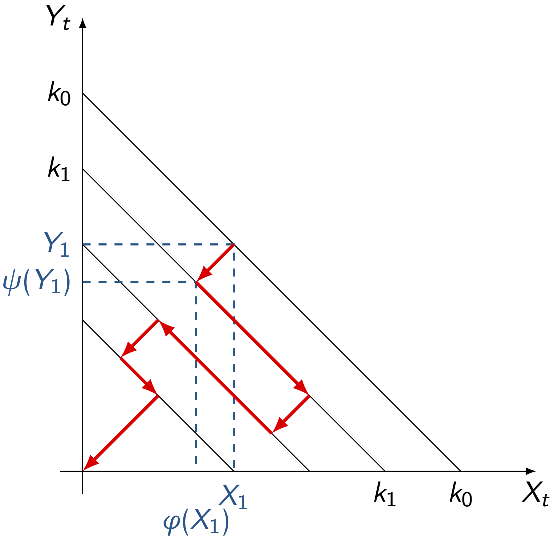
$$W_t^*(k_{t-1}) = \max_{X_t} \left\{ f(X_t) + g(k_{t-1} - X_t) + W_{t+1}^*\big(\varphi(X_t) + \psi(k_{t-1} - X_t)\big) \right\}$$

и так далее...

В момент начала управления $t = 0$:

- $k_0 = k$ — капитал еще не потрачен.
- Оптимальное значение: $W_{\max} = W_1^*(k)$
- Расходы: X_1, $y_1 = k - X_1$.

Геометрическая интерпретация



Распределение средств — движение внутрь треугольника.

Пример I

Найти оптимальный способ распределения средств $P = 100$ тыс. руб между двумя предприятиями на три года, если вложенные средства в первое предприятие дают доход $f(x) = 0.9x$ и возвращаются в размере $\varphi(x) = 0.5x$. Аналогично, для второго предприятия $g(x) = 0.8x$ и $\psi(x) = 0.7x$.

- На последнем шаге $t = 3$. Распределяемый запас k_2. Первое предприятие может получить $X_3 \in [0, k_2]$. Вычислим условный оптимум 3-го шага:

$$W_3^*(k_2) = \max_{X_3} \left\{ f(X_3) + g(k_2 - X_3) \right\} =$$

$$= \max_{X_3} \left\{ 0.9X_3 + 0.8(k_2 - X_3) \right\} = \max_{X_3 \in [0, k_2]} \left\{ 0.1X_3 + 0.8k_2 \right\} = 0.9k_2$$

При $X_3 = k_2$

Пример II

- На предпоследнем шаге $t = 2$.
  - Распределяемый запас k_1.
  - Первое предприятие может получить $X_3 \in [0, k_1]$.

Вычислим условный оптимум 2-го шага:

$$W_2^*(k_1) = \max_{X_2} \Big\{ f(X_2) + g(k_1 - X_2) + W_3^*\big(\varphi(X_2) + \psi(k_1 - X_2)\big) \Big\} =$$

$$= \max_{X_2} \Big\{ 0.9X_2 + 0.8k_1 - 0.8X_2 + 0.9\big(0.5X_2 + 0.7(k_1 - X_2)\big) \Big\} =$$

$$= \max_{X_2} \Big\{ 0.1X_2 + 0.8k_1 + 0.9(0.7k_1 - 0.2X_2) \Big\} =$$

$$= \max_{X_2 \in [0, k_1]} \{ 1.43k_1 - 0.08X_2 \} = 1.43k_1$$

При $X_2 = 0$

Пример III

- На первом шаге $t = 1$.
  - Распределяемый запас $k_0 = 100$.
  - Первое предприятие может получить $X_1 \in [0, 100]$.

Вычислим условный оптимум 2-го шага:

$$W_1^*(k_0) = \max_{X_1} \Big\{ f(X_1) + g(k_0 - X_1) + W_2^*\big(\varphi(X_1) + \psi(k_0 - X_1)\big) \Big\} =$$

$$= \max_{X_1} \Big\{ 0.9X_1 + 80 - 0.8X_1 + 1.43\big(0.5X_1 + 0.7(100 - X_1)\big) \Big\} =$$

$$= \max_{X_1} \Big\{ 0.1X_2 + 80 + 1.43(70 - 0.2X_1) \Big\} =$$

$$= \max_{X_1 \in [0,100]} \{ 180.1 - 0.186X_1 \} = 180.1$$

При $X_1 = 0$

Максимальный доход — 180.1 тыс. руб.

Пример IV

Составим оптимальный план:

Год	Ресурс	P_1	P_2
1	$k_0 = 100$	$X_1 = 0$	$Y_1 = 100$
2	$k_1 = \varphi(X_1) + \psi(Y_1) = 0.7 \cdot 100 = 70$	$X_2 = 0$	$Y_2 = 70$
3	$k_2 = \varphi(X_2) + \psi(Y_2) = 0.7 \cdot 70 = 49$	$X_3 = 49$	$Y_3 = 0$

Резюме

К настоящему моменту вы знаете:

1. Что такое метод динамического программирования.
2. Как решать задачу экономического планирования (распределения ресурсов на m лет между n предприятиями).
3.

Убедитесь, что вы не только знаете, но и умеете применять рассмотренные методы.

Задание

Вам нужно освоить:

1. Решение задачи распределения ресурсов Кремер Н. Ш. Исследование операций в экономике §12.3 с. 253–260.
2. Решение задачи о замене оборудования Кремер Н. Ш. Исследование операций в экономике §12.5 с. 265–270.

По такому поводу жду от Вас конспект этой лекции в который войдет:

- Принцип и формулы Беллмана
- Задача распределения ресурсов. Формулировка+Пример
- Задача планирования. Формулировка+Пример
- Задача о замене оборудования. Формулировка+Пример

Источники информации

- Исследование и оптимизация моделей: Кремер Н. Ш. Исследование операций в экономике глава 12 с. 245–273.
- Все материалы по курсу здесь: https://cloud.mail.ru/public/48BX/47oESuaQQ