Динамическое программирование

Динамическое программирование (или динамическое планирование) представляет собой особый математический аппарат, позволяющий осуществлять оптимальное планирование управляемых процессов. Под «управляемыми» подразумеваются процессы, на ход которых мы можем в той или другой степени влиять. (Е.С. Вентцель)

История динамического программирования

Динамическое программирование возникло и сформировалось в 1950–1953 гг. благодаря работам Ричарда Беллмана и его сотрудников. Беллман (Bellman) Ричард Эрнест (1920-84) – американский математик. Первыми задачами, решаемыми с помощью динамического программирования были задачи управления запасами.

Задача динамического программирования

Задачи динамического программирования укладываются в следующую схему:

I. Имеется набор способов действий – допустимых управлений.

II. Имеется целевая функция («прибыль», «убыток») $S = S(u)$, где u пробегает допустимые управления.

Требуется выбрать управление, которому отвечает оптимальное значение целевой функции.

Формализация задачи динамического программирования

Динамическая система (ДС) – такой объект, который может развиваться.

Квазидинамическая система – система, которую можно привести к динамической.

Пространство состояний (фазовое пространство) динамической системы – множество всех возможных состояний динамической системы.

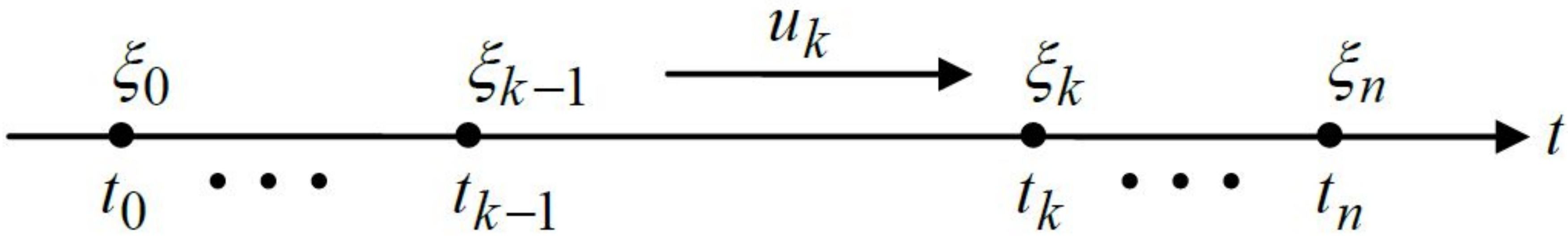
ДС с дискретным временем – ДС, состояние которой меняется в дискретные моменты времени, например, t_0, $t_1, t_2, .., t_n$.

Траектория – набор состояний, через которые проходит ДС от начального к конечному состоянию.

Управляемая ДС – ДС, траекторию которой можно менять с помощью управляющих импульсов U_k.

Функционирование динамической модели

1. В начальный момент времени t_0 система находится в фиксированном состоянии ξ_0.
2. Переход $\xi_{k-1} \rightarrow \xi_k$ от состояния в момент t_{k-1} к состоянию в момент t_k (от t_0 к t_1 от t_1 к t_2 и т.д.) осуществляется под воздействие управляющих сигналов u_k, т.е. $\xi_k = \phi(\xi_{k-1}, u_k)$.
Набор состояний $\xi_0, \xi_1, \ldots, \xi_n$ – траектория ДС. Причем, начало всех траекторий одинаковое - ξ_0.



Общая задача динамического программирования

Пусть имеется управляемая динамическая система оцениваемся каким-либо показателем, например, суммарным доходом $S=S(u_1,u_2,..,u_n)$. Суммарный доход равен сумме доходов на каждом шаге

$$S=f_1+f_2+\ldots+f_n. \quad (1)$$

f_k – доход на k-ом шаге, который зависит от состояния ДС в начале k-го шага и выбранного управления uk:

$$f_k = f_k(\xi_{k-1}, u_k) \quad (2)$$

Подставляя (1) в (2) получим:

$$S = \sum_{k=1}^{n} f_k(\xi_{k-1}, u_k)$$

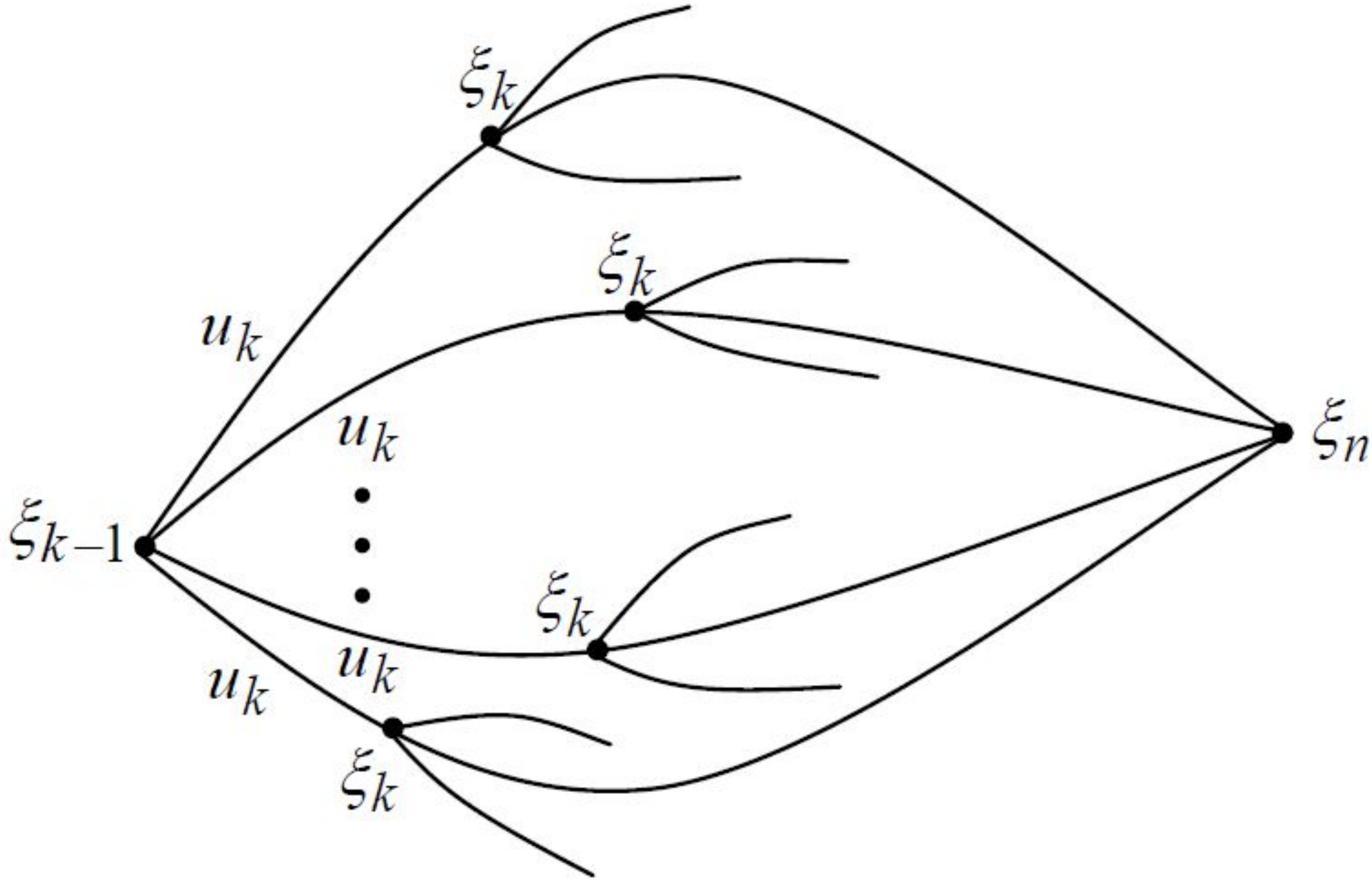
Такая функция называется **аддитивной целевой функцией**.

Общая задача динамического программирования

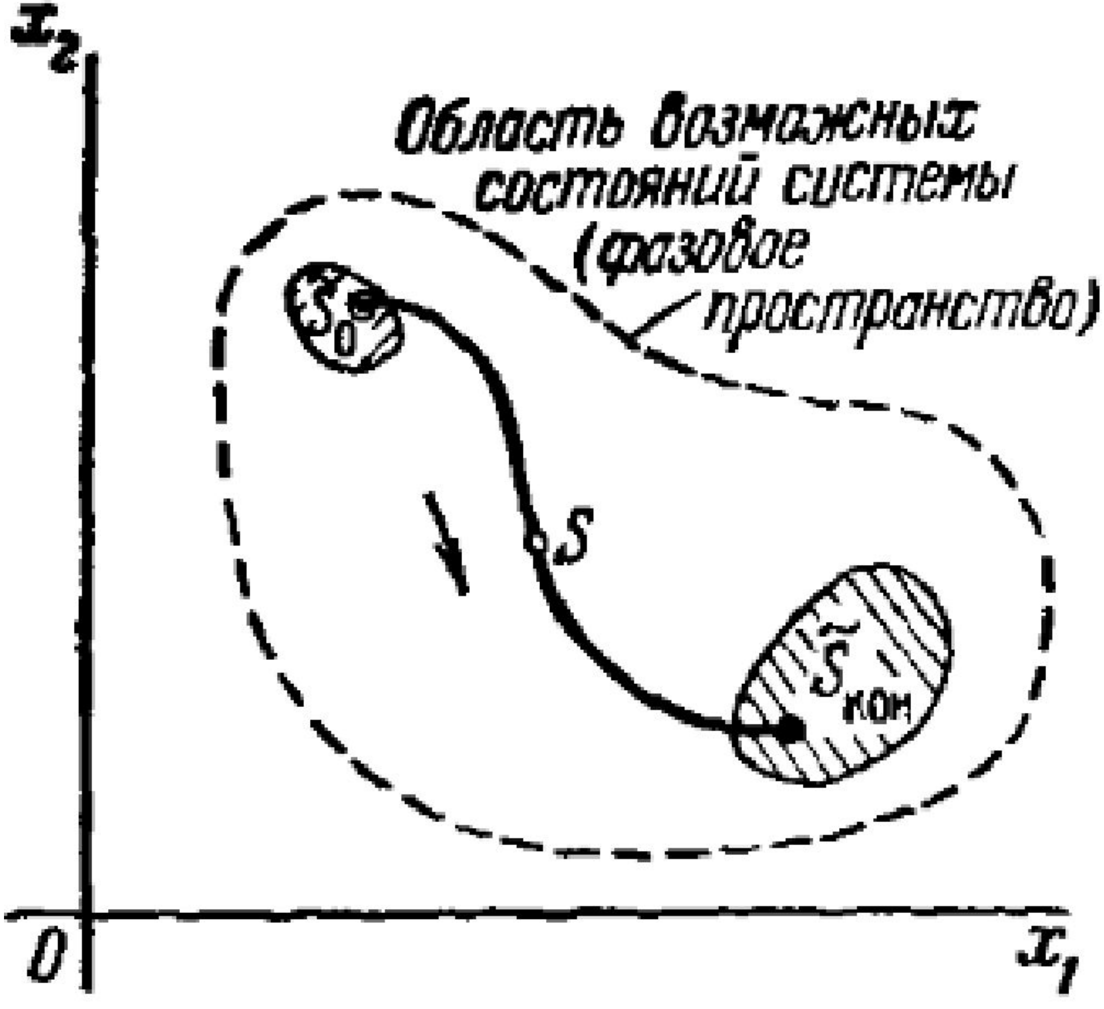
Для анализа имеется управляемая динамическая система, для которой выделено n шагов, а на каждом шаге выделены все возможные состояния (ξ_k), через которые проходит система, а также выделено начальное состояние ξ_0, а на каждом шаге заданы возможные управляющие воздействия (u_k). Также имеется аддитивная функция.

Задача ДП состоит в том, чтобы найти такую последовательность управляющих воздействий во время функционирования ДС ($u_1, u_2, \ldots, u_n$), чтобы аддитивная функция S достигала максимума или минимума. Траектория, на которой достигается max или min целевой функции, называется **оптимальной траекторией**. Задачей ДП является отыскание оптимальной траектории.

Метод решения задачи динамического программирования



Метод решения задачи динамического программирования



Метод решения задачи динамического программирования

Предположим, что к на- чалу *k*-го шага система оказалась в состоянии ξ_{k-1}. Оставшийся путь до конца система может проходить по различным траекториям в зависимости от выбора последующих управлений. Каждой траектории отвечает свой суммарный доход, т.е. доход на участке Обозначим этот доход через S_k.

$$S_k=f_k+f_{k+1}+\ldots+f_n$$

$S^*_k(\xi_{k-1})$ – **условный максимальный доход** – максимальный суммарный доход, начиная с *k*-го шага и до конца, т.е. на участке [k-1,n]. Тогда основную задачу ДП можно формал

$$S_k^*(\xi_{k-1}) = \max_{u_k}\{f_k(\xi_{k-1},u_k)+S_{k+1}^*(\xi_k)\}$$

И требуется найти условное оптимальное управление на k-м шаге $u^*_k(\xi_{k-1})$.

Метод решения задачи динамического программирования

Задачи ДП решаются в два этапа:

1. движение от конца к началу

Начиная с кон- ца последовательно находим:

$$S_n^*(\xi_{n-1}), u_n^*(\xi_{n-1}); S_{n-1}^*(\xi_{n-2}), u_{n-1}^*(\xi_{n-2}); \ldots; S_1^*(\xi_0), u_1^*(\xi_0)$$

для всех возможных на соответствующих шагах состояний с использованием на каждом шаге, начиная с *n-1*-го, формулы для

$$S_k^*(\xi_{k-1}) = \max_{u_k}\{f_k(\xi_{k-1}, u_k) + S_{k+1}^*(\xi_k)\}$$

Метод решения задачи динамического программирования

Задачи ДП решаются в два этапа:

1. движение от конца к началу

Начиная с кон- ца последовательно находим:

$$S_n^*(\xi_{n-1}), u_n^*(\xi_{n-1}); S_{n-1}^*(\xi_{n-2}), u_{n-1}^*(\xi_{n-2}); \ldots; S_1^*(\xi_0), u_1^*(\xi_0)$$

для всех возможных на соответствующих шагах состояний с использованием на каждом шаге, начиная с *n-1*-го, формулы для

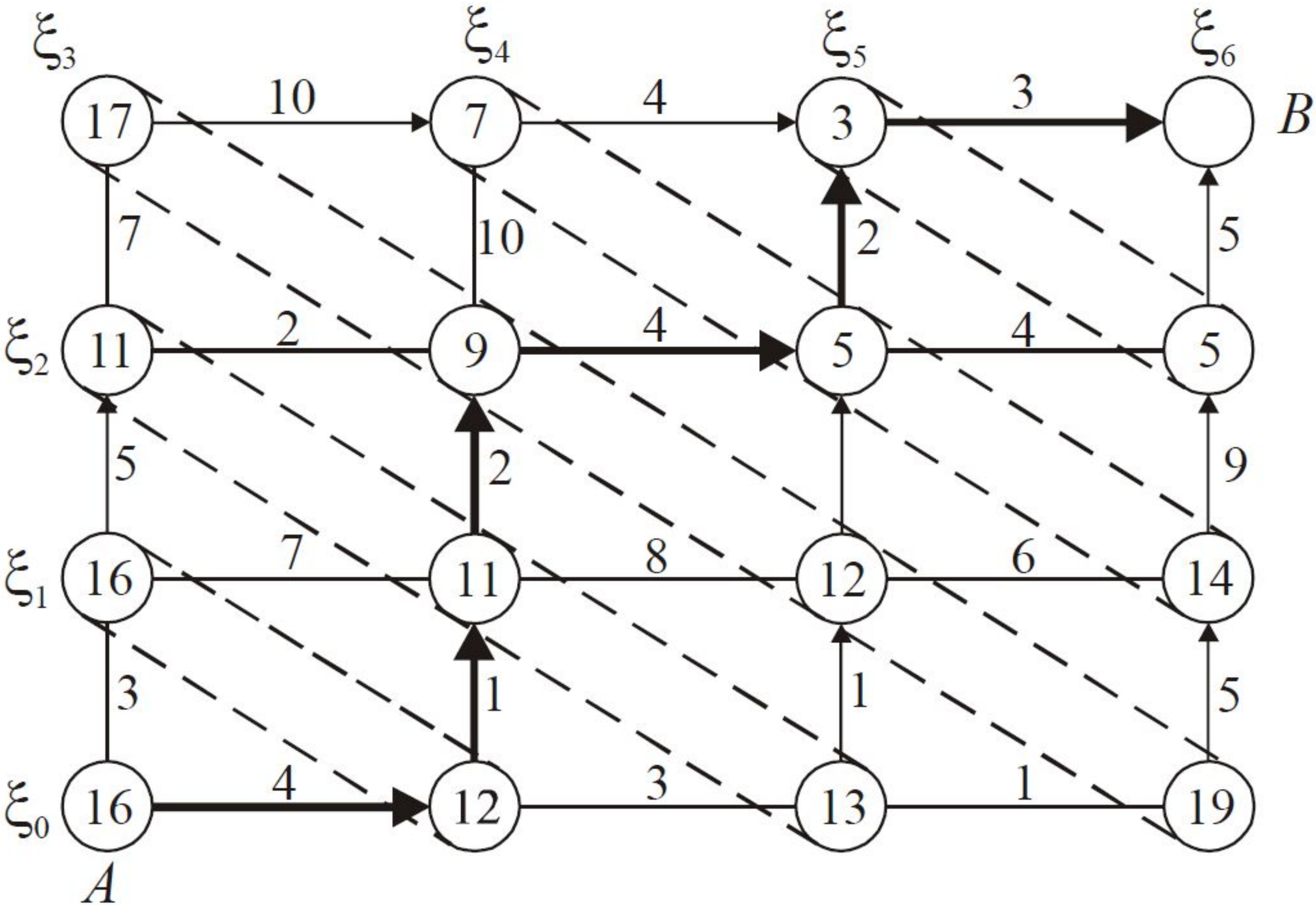
$$S_n^*(\xi_{n-1})$$

2 движение от начала к концу

Двигаясь от начала, существенно используя закрепленность начального состояния, строим безусловную оптимальную т

$$\xi_0 \xrightarrow{u_1^*(\xi_0)} \xi_1^* \xrightarrow{u_2^*(\xi_1^*)} \xi_2^* \longrightarrow \ldots \longrightarrow \xi_n^*$$

Задача об оптимальном маршруте



Задача об оптимальном маршруте

Требуется найти оптимальный маршрут на плоскости, проходящий через некоторые точки.

ДС представляет собой граф, где каждая вершина имеем четырех соседей. В нашем случае потребуется 6 шагов.

Состояние – узел графа. A – начальное состояние B – конечное состояние.

Под управлением будем понимать выбор направления движения: по горизонтали или вертикали.

Аддитивная целевая функция – величина потерь, приписанная к каждой дуге графа.

Целевая функция – суммарные потери на всех 6 шагах.

Решение:

Помечаем узлы графа последовательно от конца к началу минимальным весом, затем от начала к концу выбираем маршрут, с наименьшим весом для каждого шага.

Задача об оптимальной последовательности погрузки и разгрузки

Пусть на оптовую базу прибыло n машин с товаром для разгрузки и m машин для загрузки товаров, направляемых в магазины. Материально ответственное лицо оптовой базы осуществляет оформление документов по операциям разгрузки или загрузки одной машины, а затем переходит к обслуживанию другой машины. Издержки от операций обусловлены простоем транспорта, типом операции (прием или отгрузка товара). Необходимо спланировать последовательность операций обоих видов таким образом, чтобы суммарные издержки по приему и отправке товаров для всех машин были минимальными.

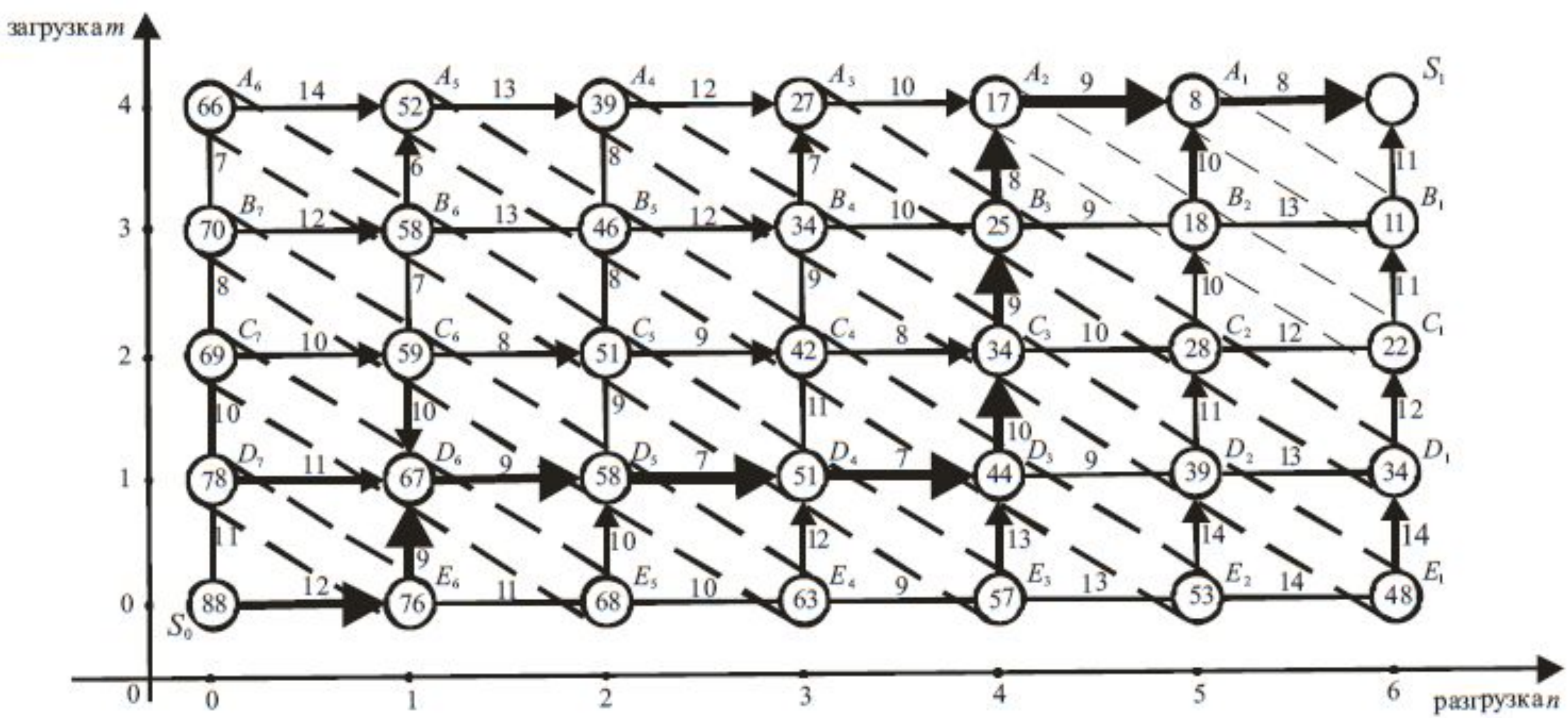
Задача об оптимальной последовательности погрузки и разгрузки

Если по оси X отложить число n разгруженных машин, а по оси Y число m загруженных товаром машин, то можно построить на плоскости граф состояний процесса, в котором каждая вершина характеризует состояние операции приема и отгрузки то- вара на оптовой базе. Ребра означают выполнение работы по приему или отправке товара на очередной машине. Каждому ребру можно сопоставить издержки, связанные с выполнением операции по разгрузке или загрузке машины.

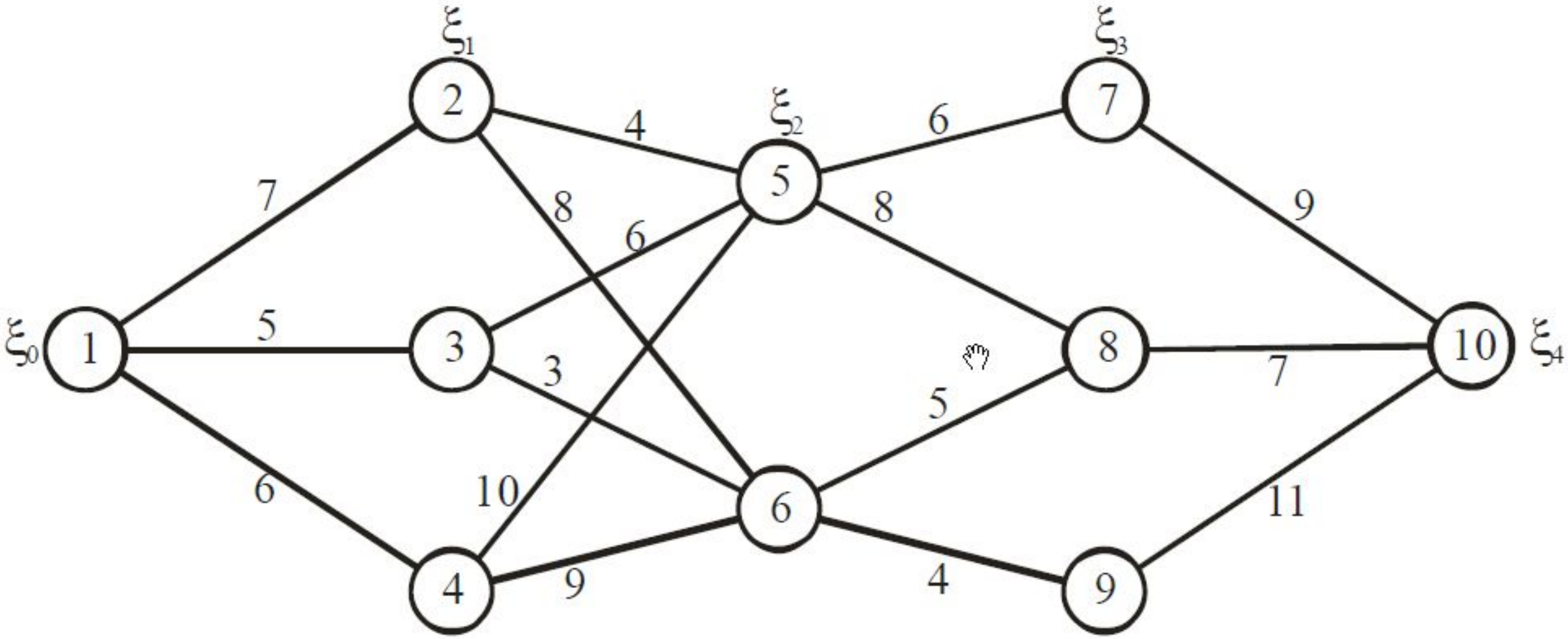
Таким образом, задачи сводится к ранее рассмотренной задаче выбора оптимального маршрута.

Задача об оптимальной последовательности погрузки и разгрузки

Пример. Пусть n=4, m=6 известны затраты по выполнению каждой операции, которые показаны на ребрах графа (рис. 4). Точка определяет начало процесса, точка конечное состояние, соответствующее приему и отправке всех машин.

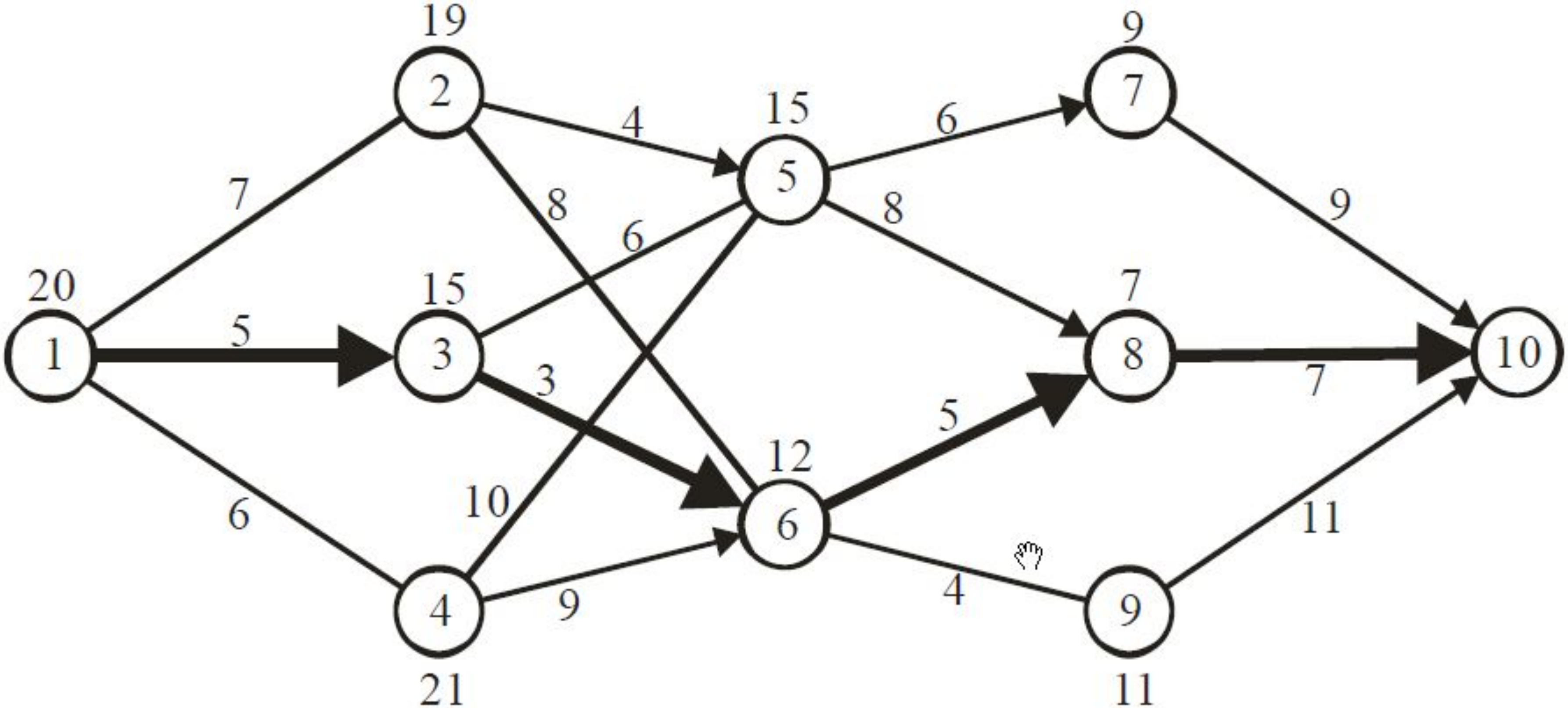


Задача о выборе оптимального пути на графе

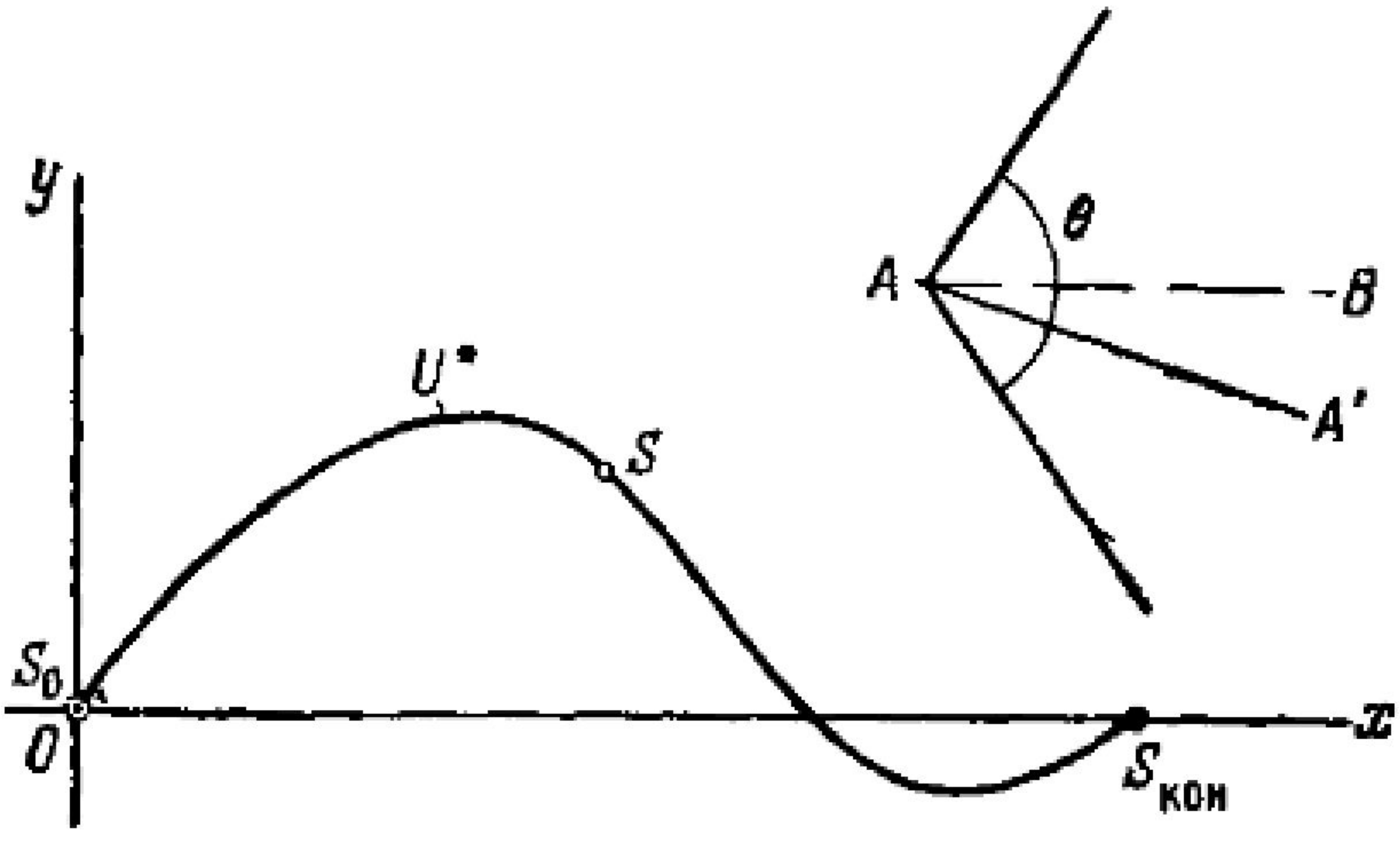


Необходимо выбрать путь из пункта 1 в пункт 10. Двигаться можно только слева направо

Задача о выборе оптимального пути на графе



Задача прокладки непрерывного оптимального пути



Задача прокладки непрерывного оптимального пути

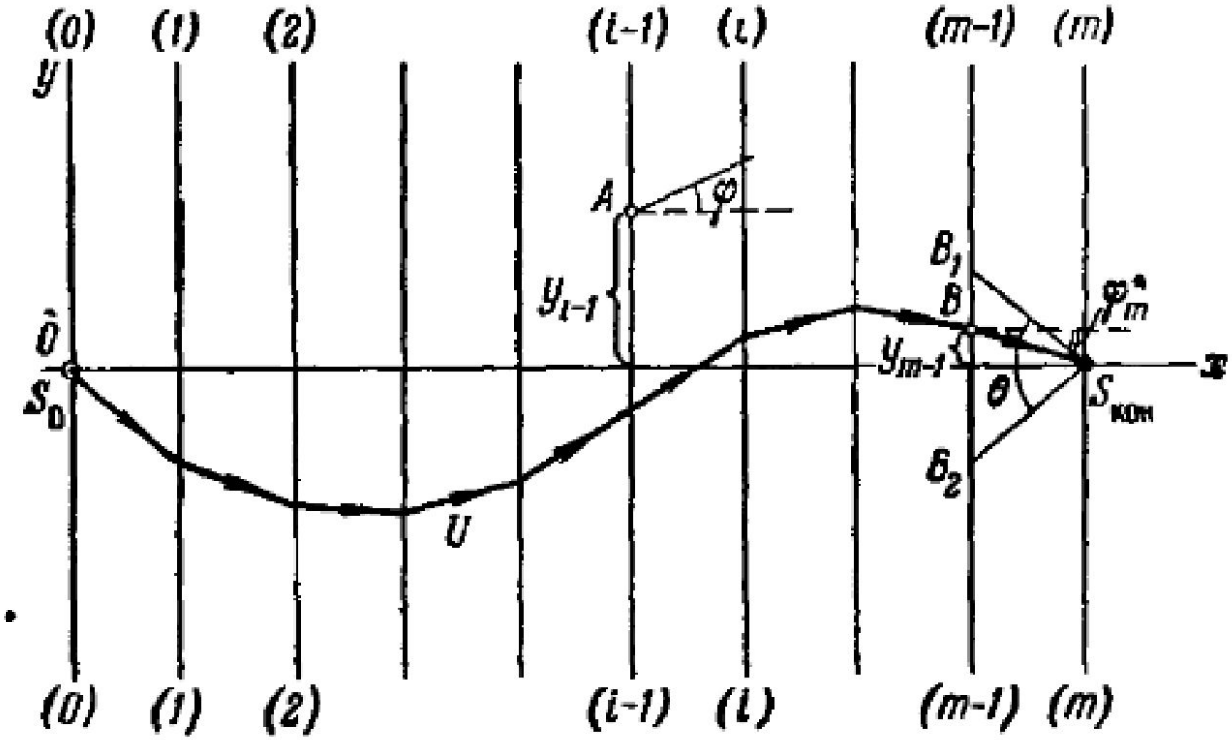
$$U_i^*(y_{i-1}) = \varphi_i^*(y_{i-1})$$

$$\varphi_m^* = \varphi_m^*(y_{m-1})$$

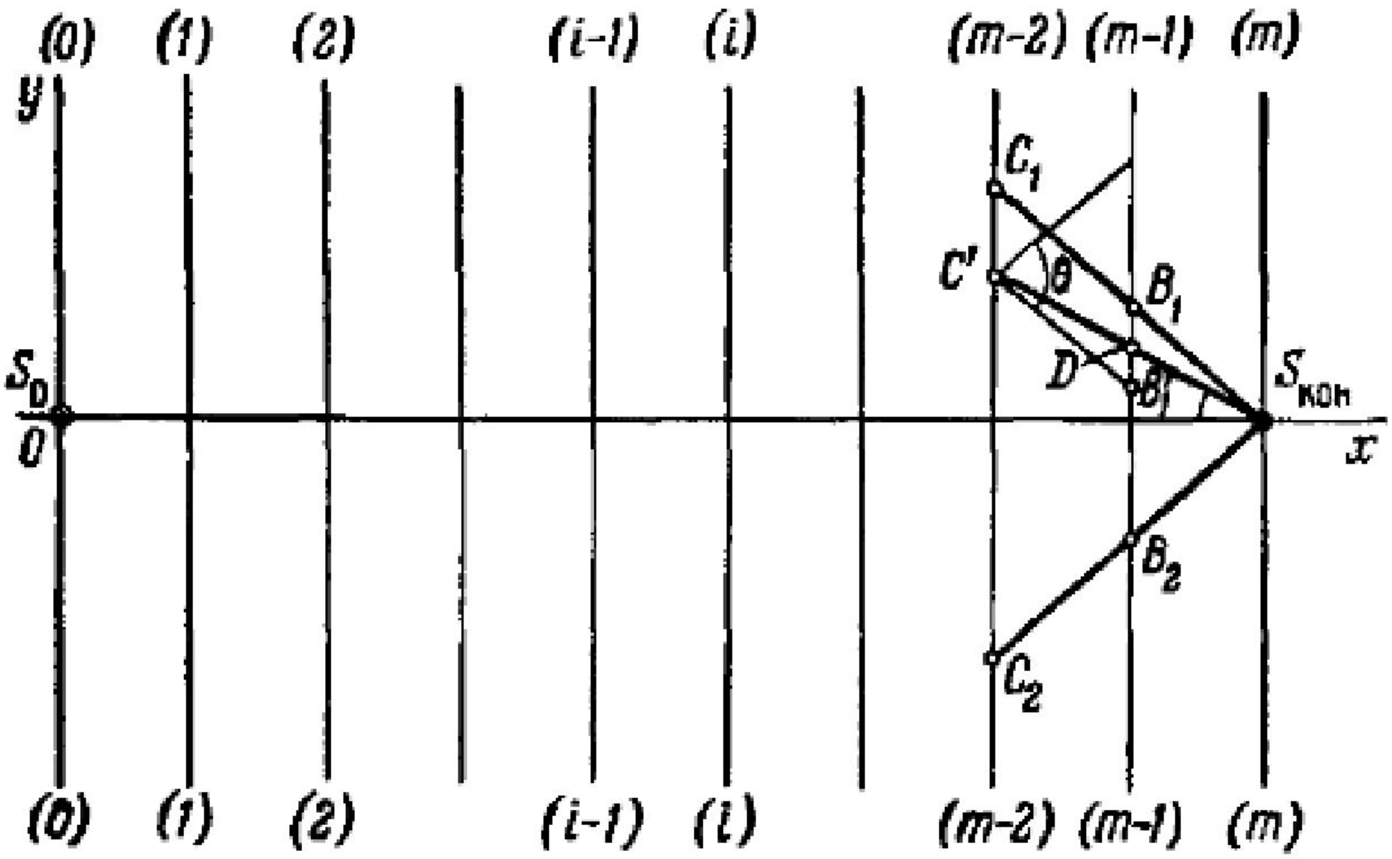
$$T_m^* = T_m^*(y_{m-1})$$

$$\varphi_m^*(y_{m-1}^{(1)}); \quad \varphi_m^*(y_{m-1}^{(2)}); \quad \ldots$$

$$T_m^*(y_{m-1}^{(1)}); \quad T_m^*(y_{m-1}^{(2)}); \quad \ldots$$

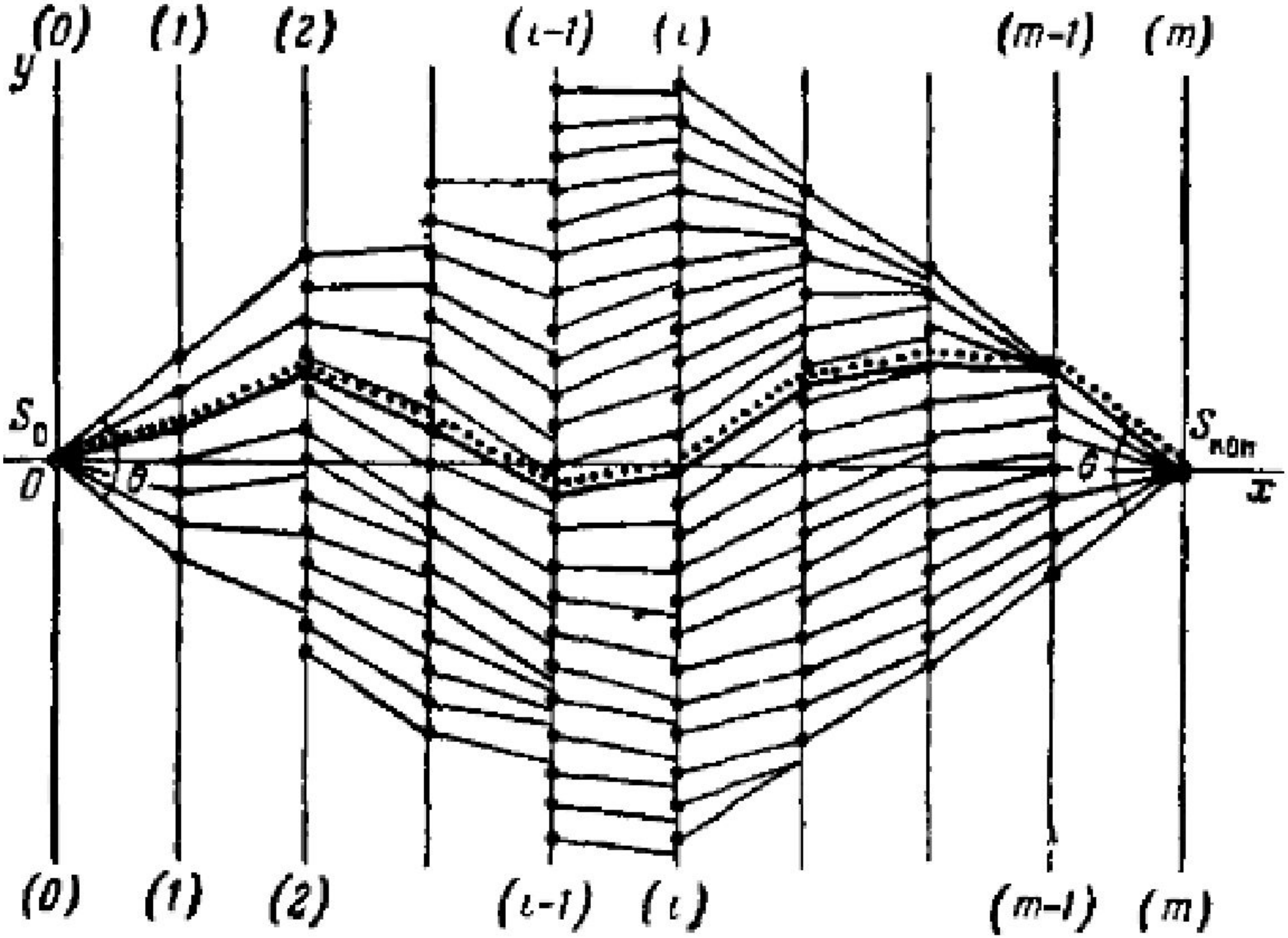


Задача прокладки непрерывного оптимального

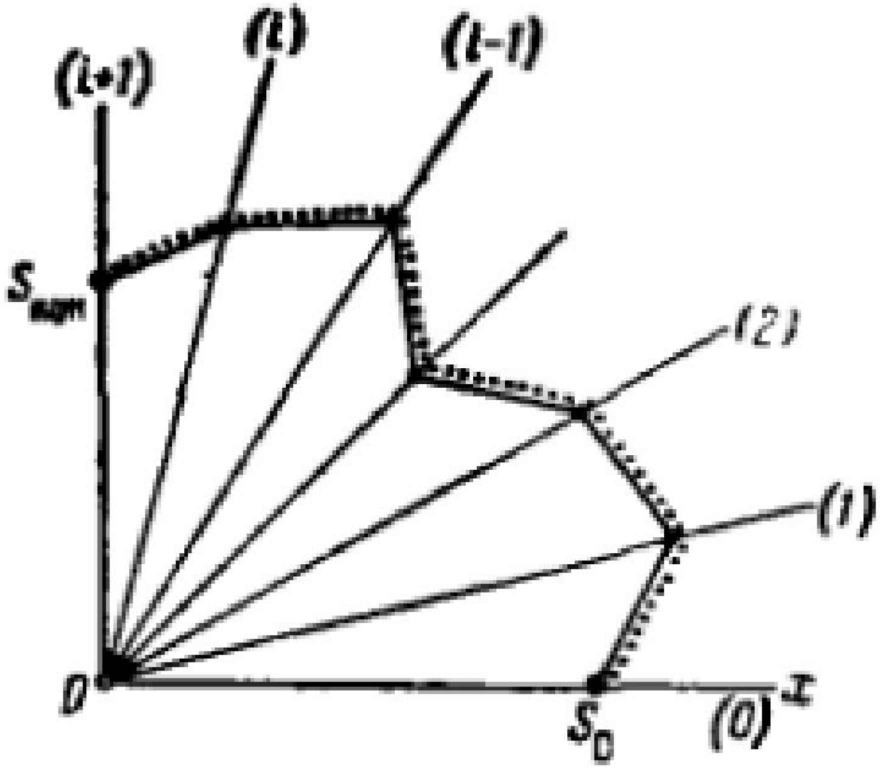


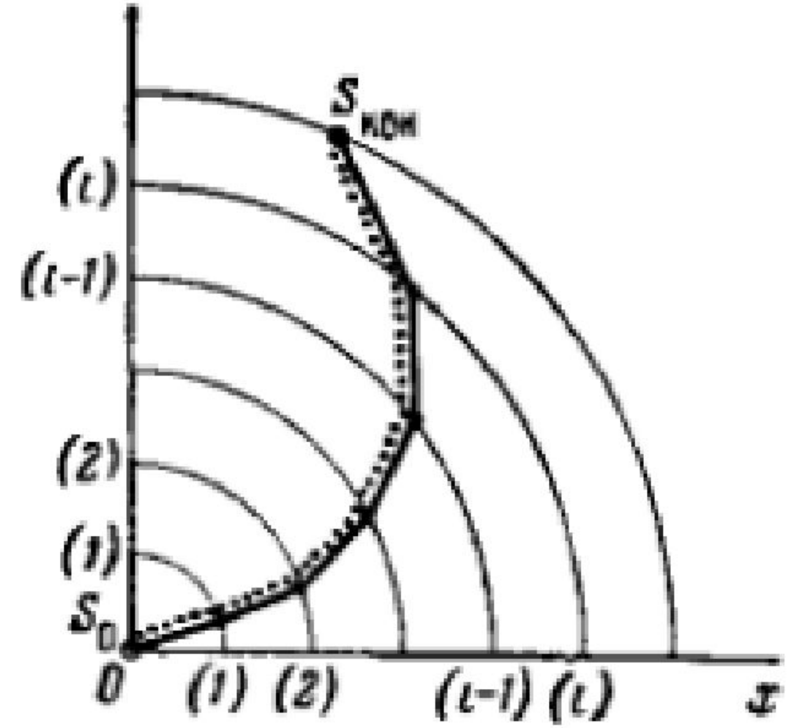
Зафиксируем на оси абсцисс несколько точек, в которых вычислим аддитивную целевую функцию.

Задача прокладки непрерывного оптимального пути

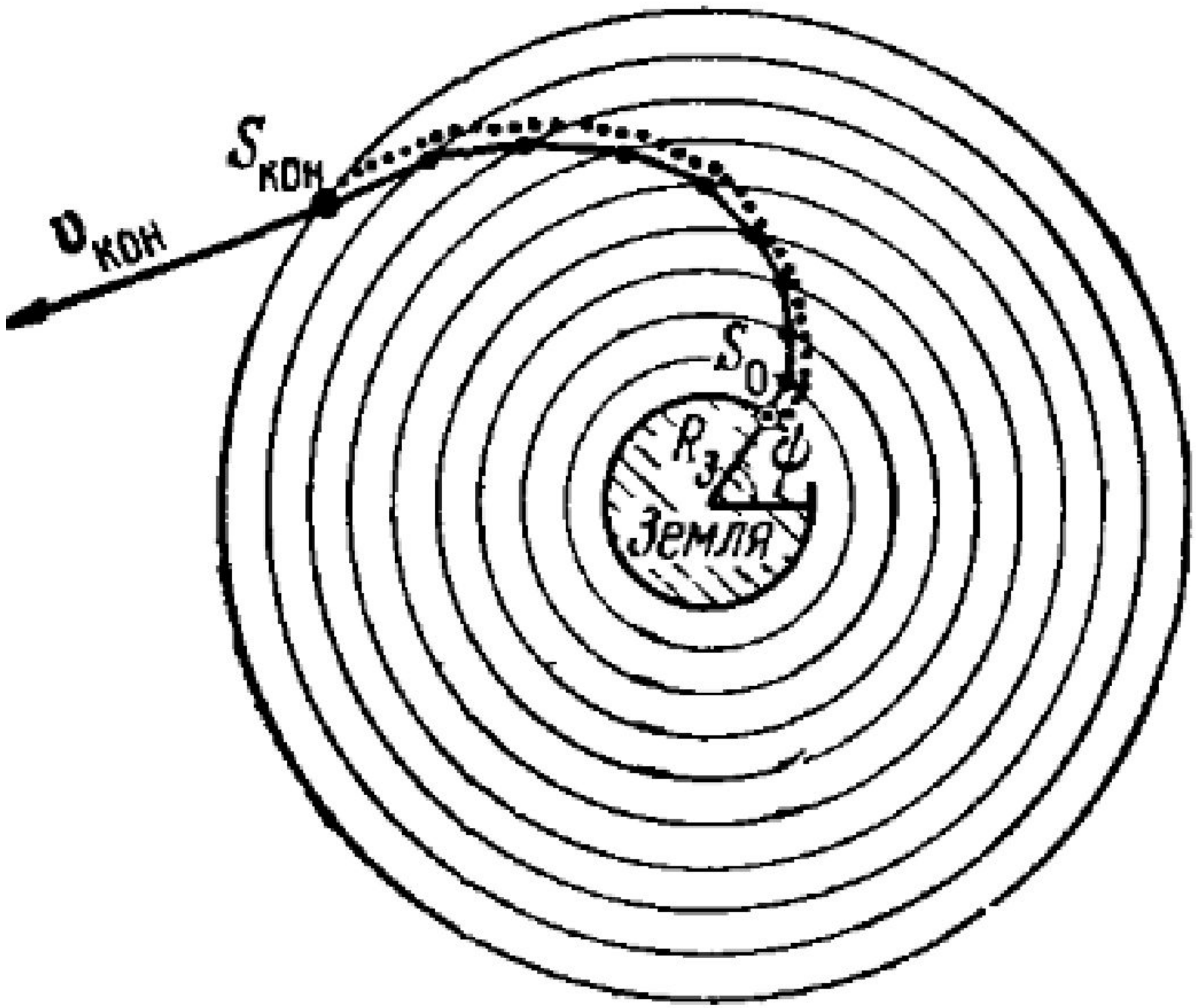


Задача прокладки непрерывного оптимального пути в полярной системе координат





Задача прокладки непрерывного оптимального пути в полярной системе координат



Задача о выборе оптимального оптимальной стратегии замены оборудования

Определить оптимальные сроки замены оборудования в течение n лет, при которых прибыль от эксплуатации оборудования максимальна, если известны: p – начальная стоимость оборудования; $R(t)$ – стоимость производимой продукции на оборудовании возраста t лет; r(t) – ежегодные затраты на эксплуатацию оборудования возраста t лет; $\phi(t)$ – ликвидная стоимость оборудования возраста t лет. Предполагаются, что к началу планового периода оборудование является новым.

Задача о выборе оптимального оптимальной стратегии замены оборудования

В качестве управления будем использовать решение о замене или незамене оборудования:

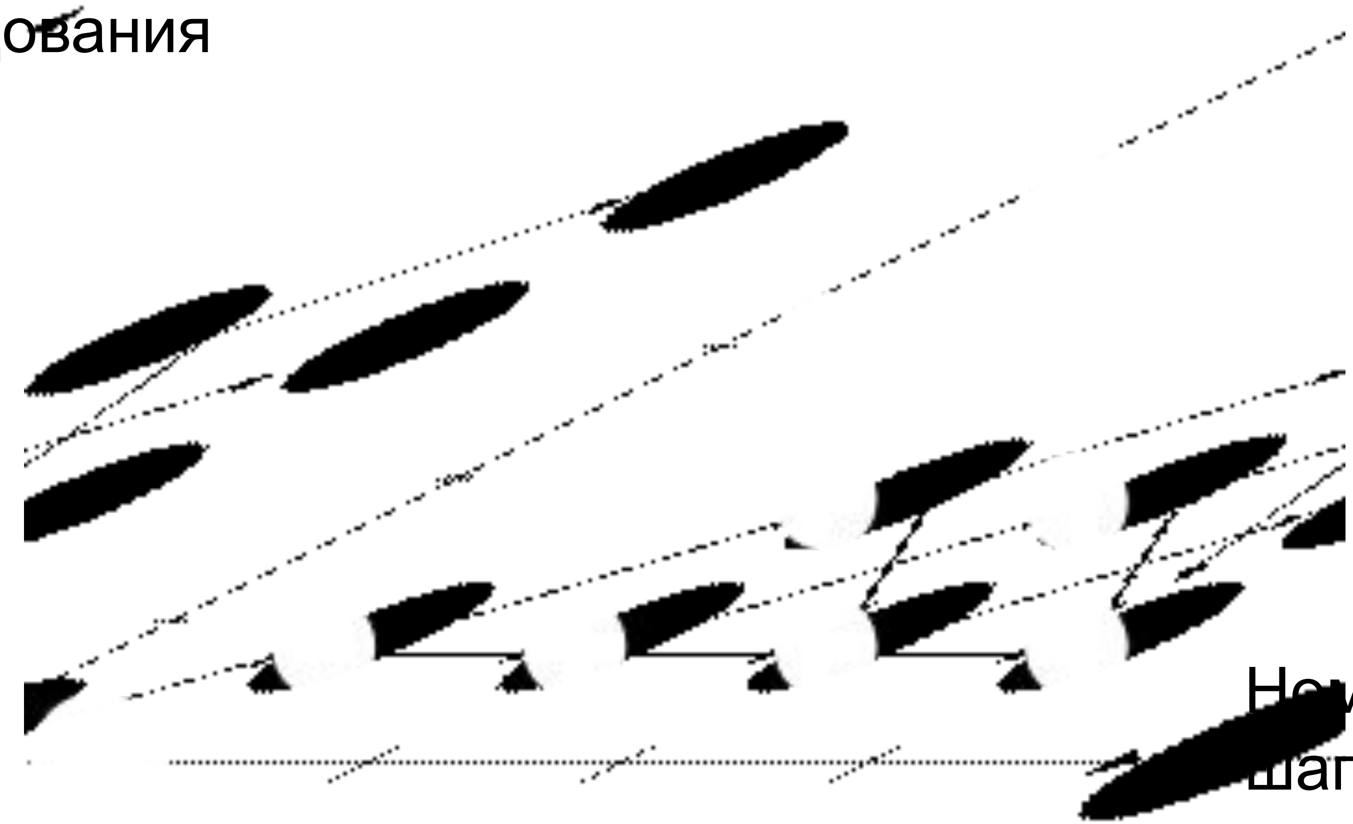
$$\xi_k = \begin{cases} \xi_{k-1} + 1, & u_k = u \\ 1, & u_k = \bar{u}. \end{cases}$$

$$f_k(\xi_{k-1}, u_k) = \begin{cases} R(\xi_{k-1}) - r(\xi_{k-1}) & при \quad u_k = u \\ R(0) - r(0) + \varphi(\xi_{k-1}) - p & при \quad u_k = \bar{u}. \end{cases}$$

$$S = \sum_{k=1}^{n} f_k(\xi_{k-1}, u_k)$$

Геометрическая интерпретация задачи о выборе оптимальной стратегии замены оборудования

Возраст оборудования



Номер шага

Задача о выборе оптимального оптимальной стратегии замены оборудования

	0	1	2	3	4	5
$R(t)$	80	75	65	60	60	55
$\mathrm{r}(t)$	20	25	30	35	45	55

Замена оборудования p=40 у.
е.

$$f_k(\xi_{k-1}, u_k) = \begin{cases} R(\xi_{k-1}) - \mathrm{r}(\xi_{k-1}) & при \quad u_k = u \\ 80 - 20 - 40 = 20 & при \quad u_k = \overline{u}. \end{cases}$$

Задача о выборе оптимального оптимальной стратегии замены оборудования

ξ_k	$S_5^*(\xi_4)$	$u_5^*(\xi_4)$	$S_4^*(\xi_3)$	$u_4^*(\xi_3)$	$S_3^*(\xi_2)$	$u_3^*(\xi_2)$	$S_2^*(\xi_1)$	$u_2^*(\xi_1)$	$S_1^*(\xi_0)$	$u_1^*(\xi_0)$
0	60	u	110	u	145	u	180	u	215	u
1	50	u	85	u	120	u	155	u		
2	35	u	70	$\overline{u}$	105	$u, \overline{u}$				
3	25	u	70	$\overline{u}$						
4	20	$\overline{u}$								
5										

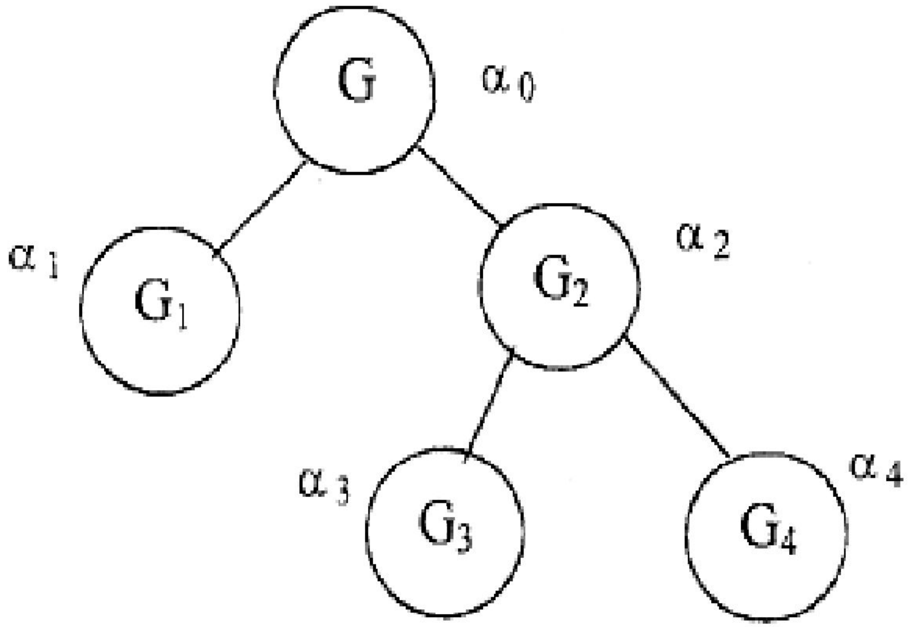
Оптимальные стратегии:

u u u $\overline{u}$ u

u u $\overline{u}$ u u

Метод ветвей и границ

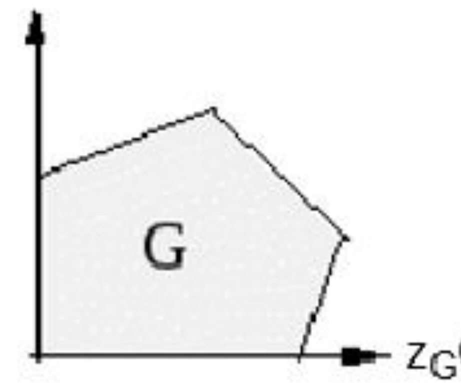
Этот метод представляет собой упорядоченный перебор вариантов и рассмотрение лишь тех из них, которые по некоторым признакам оказываются перспективными.



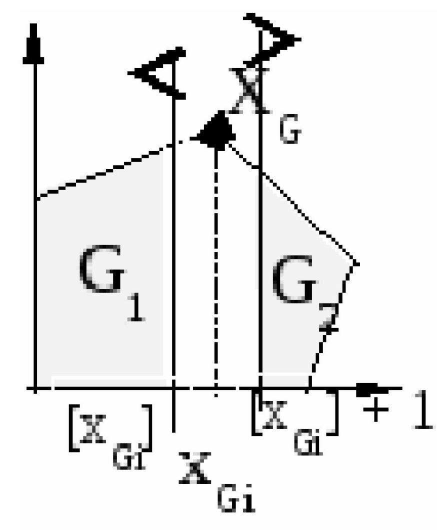
Метод ветвей и границ в задаче целочисленного программирования

Решив задачу линейного программирования получим область допустимых планов (ОДП),

$$D \begin{cases} \left.\begin{array}{l} AX = B \\ X \geq 0 \end{array}\right\} G \\ X \in Z \end{cases}$$



Исходное множество – G.
Если план Z_G – целочисленный, то задача решена.
Иначе выбираем один из нецелочисленных параметров и выбираем к нему две ближайших целых числа.
D2={X_D: $x_i[X_{Gi}] + 1$}, где [] – целая часть числа.



Выбираем наиболее оптимальное

Пример применения метода ветвей и границ

Задача ЛП:

$\max 2x_1 + x_2$

$5x_1 + 2x_2 10$

$3x_1 + 8x_2 13$

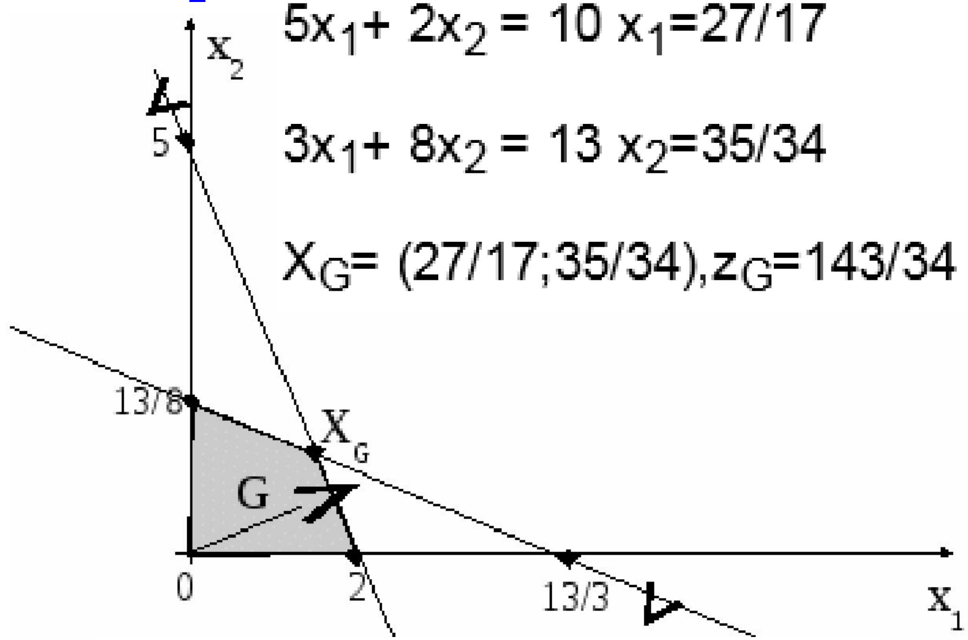
$x_{1,2} 0$

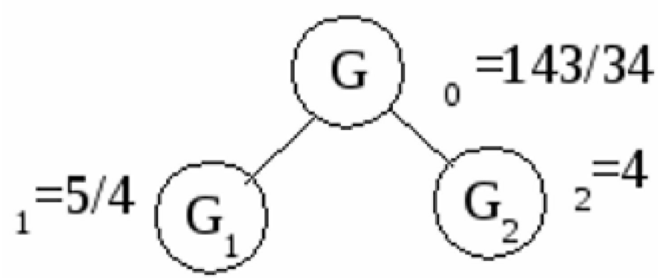
$x_{1,2} Z$

$5x_1 + 2x_2 = 10$ $x_1 = 27/17$

$3x_1 + 8x_2 = 13$ $x_2 = 35/34$

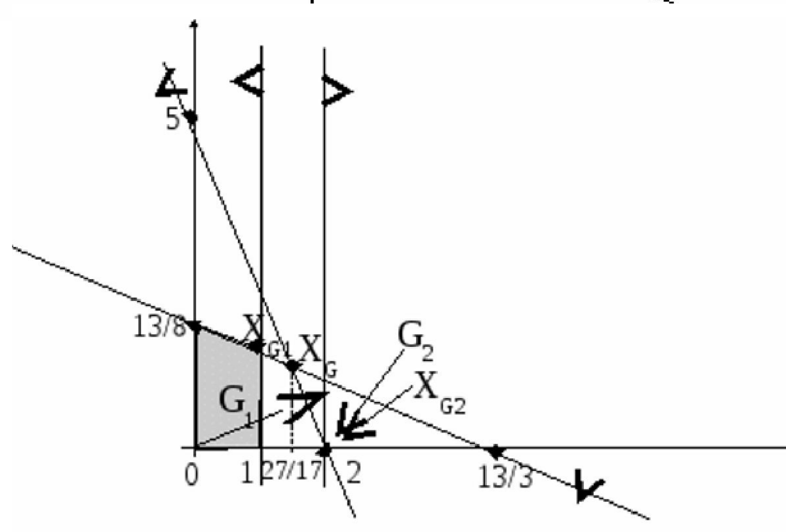
$X_G = (27/17; 35/34), z_G = 143/34$





$G_1 = \{X G: x_1 1\}$

$G_2 = \{X G: x_1 2\}$



Пример применения метода ветвей и границ

Проверим, не является ли точка G_1 оптимальной, для этого решим систему уравнений

$$\begin{cases} x_1 = 1 & x_1=1 \\ 3x_1 + 8x_2 = 13 & x_2=5/4 \end{cases}$$

$X_{G1} = (1;\ 5/4), z_G=13/4$

G1<G2 => G2 – оптимальная точка. $X^*=X_{G2}=(2;0)$, $z^*=4$

Задача распределения ресурсов

Объёмы	10	20	30	40	50
Подразделение 1	12	28	32	42	58
Подразделение 2	15	26	34	41	52
Подразделение 3	11	23	-	45	56
Подразделение 4	16	25	33	41	53

Необходимо вложить 80 у.е., чтобы получить максимальный доход

Выделяемые объёмы / Подразделения	10	20	30	40	50
1	12	28	32	42	58

1 / 2	10-12	20-28*	30-32	40-42	50-58
10–15*	20–27	30-43*	40-47	50-57	60–73*
20–26	30-38	40-54*	50-58	60-68	70–84*
30-34	40-46	50-62*	60-66	70-76	80–92*
40-41	50-53	60-69	70-73	80–83	
50-52	60-64	70-80	80-84		

Задача распределения ресурсов

3 \ 1+2	10-15	20-28*	30-43*	40-54*	50-62	60-73	70-78	80–92
10–11	20–26	30-39	40-54*	50-65	60–73	70–84	80–95	
20–23	30-38	40-51	50-66*	60-77*	70–85	80-96		
40-45	50-60	60-73	70-88*	80–99*				
50-56	60-71	70-84	80-99					

4 \ 1+2+3	30-43	40-54	50-66	60-77	70-88	80–99
10–16					80–104*	
20–25				80-102		
30-33			80-99			
40-41		80–95				
50-53	80-96					

Литература

1. Романовская, Адель Матвеевна Динамическое программирование: Учебное пособие. Романовская А.М., Мендзив М.В.– Омск: Издатель Омский институт (филиал) РГТЭУ, 2010. – 58 с.

2. Е.С. Вентцель Элементы динамического программирования Издательство «Наука», Москва, 1961