КУРС «МЕТОДЫ ОПТИМИЗАЦИИ»

ЧАСТЬ 1

Методы оптимизации, их место в теории исследования операций.

Метод оптимизации как дисциплина, представляет собой раздел теории исследования операций, задачей которой является поиск в рамках принятой модели решений, отвечающих экстремальным значениям критерия.
Операцией называется совокупность взаимосогласованных действий направленных на достижение определенной цели. До тех пор, пока цель не определена, нет смысла говорить об операции. Если же цель определена, как правило, есть несколько путей ее достижения, среди которых необходимо найти лучший. Понятие *лучший* что-либо означает, когда определен показатель или критерий качества выбираемых решений.
Любую операцию можно охарактеризовать следующими составляющими:

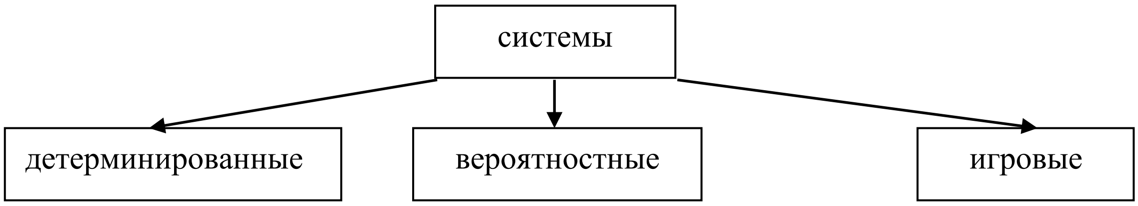
- *Стратегией* (вариант решения) называется допустимый или возможный способ достижения цели.
- *Действующими факторами операции* называются объективные условия, определяющие ее особенности и влияющие на ее исход.
- *Критерием эффективности* операции или стратегии называется показатель достигнутого соответствия между полученным результатом и целью операции.
- *Состоянием операции* в некоторый момент времени называют совокупность ее характеристик проявляющихся в данный момент.

Под *параметром стратегии* понимается локальное качество операции, обусловливаемое его действующими факторами и учитываемое критерием эффективности стратегии.
Под *моделированием параметров* (или соответственно моделированием системы) понимается определение зависимостей изменения локальных качеств (или критериев оптимальности) с учетом действующих факторов операции.
Математической моделью операции называются формальные соотношения, устанавливающие связь принятого критерия эффективности с действующими факторами операции.
Решением (ограниченным множеством решений) связанным с выбранной математической моделью называется конкретный набор значений параметров полученных в результате использования этой математической модели.
Схематично этапы исследования операций можно представить следующим образом:

Содержание этапов	*этапы*
Формулировка цели операции, обоснованный выбор критерия эффективности	1
Анализ системы с целью: 1. выбор исходного множества стратегий. 2. обоснованный выбор множеством исходных параметров, описывающих стратегию с учетом сформулированной цели операции.	2
При необходимости моделировании параметров стратегии (построение математической модели параметров)	3
Построение математической модели операции – получение формальной зависимости, изменения критерия качества от изменения параметров системы.	4
Получение решения – выбор оптимального набора параметров (соответственно оптимальной стратегии с учетом сформулированной цели операции).	5

Место методов оптимизации в общей теории исследования операций – оптимальным образом решение формализованной задачи. Методы оптимизации используются на 4-5, при необходимости на 3.

Системой называется упорядоченная совокупность материальных объектов (элементов) объединенных какими-либо связями (напр. механич, электрич и т.д.) предназначенных для достижения определенной цели.
Если исходить из понятия, что поведение системы есть последовательный ряд операций, то системы можно классифицировать следующим образом:



Детерминированной считается такая система, в которой составные части взаимодействуют друг с другом точно предсказуемым образом. Ее поведение предсказуемо, если известно текущее состояние элементов и законы преобразования циркулирующие между ними.
Вероятностной называют систему, возможное поведение которой и его последствия описываются на языке теории вероятности, здесь знание текущего состояния и особенности взаимной связи элементов не достаточно для однозначного предсказания будущего, что обуславливает необходимость вероятностной оценки.
Игровой является система, осуществляющая разумный выбор своего поведения в будущем. В основе выбора лежат неформальные соображения, руководствоваться которыми может лишь человек.
Другим важным классификационным признаком системы является ее степень сложности. По этому признаку их разделяют на:

- *простые* состояния, которые не многочисленны и легко поддаются описанию.
- *сложные* отличающиеся разнообразием внутренних связей, но допускающие их описание.
- *крупномасштабные*, не поддающиеся какому-либо описанию связей.

Важное место в исследованиях различного характера является структура системы, и возможность ее *расчленяемости* на подсистемы, каждая из которых может иметь свой критерий эффективности.
Принцип расчленяемости системы находится, в основе так называемого *системного подхода* состоящего в том, что задачу оптимизации можно упростить, сформулировав и решив ее для отдельных подсистем с учетом того, в какой мере критерий эффективности подсистемы влияет на критерий эффективности системы в целом.
Эвристика оказывается полезной и независимой при решении задач имеющих не числовую природу или характеризуемых сложностью определения каких-нибудь параметров либо их формализованного сравнения.
Методы оптимизации классифицируются по числу критериев оптимальности. Их подразделяют на *однокритериальные* либо *многокритериальные*. В последнем случае при выборе оптимального решения необходимо учитывать уже не один критерий, а их совокупность, как правило, характеризуемых разнородными параметрами, что усложняет их сравнительную оценку. Здесь, как правило, не обойтись без эвристики.
В рамках задач однокритериальной оптимизации наибольший интерес представляют методы, относящиеся к группе оптимального планирования, при ограниченных ресурсах. Данными методами учитываются разнообразные задачи синтеза расписаний и создания новой техники. С формальной точки зрения данная задача относится к задачам математического программирования (термин программирование, здесь понимается в смысле составления плана, либо программы действий).
Задача оптимизации состоит в выборе наилучших планов.

Формальная постановка задачи математического программирования (М П).

Не смотря на смысловое разнообразие задач, все они формально сводятся к одной общей постановке.
Найти значение переменных обращающих в максимум, либо минимум значение целевой функции:

$$\min(\max) \leftarrow Z = f(x_{1,\ldots,} x_n),$$

при условиях: $g_i(x_{1,\ldots,} x_n) \lesseqgtr bi,\ i=1,m$
Замечание: в случае, когда Z есть множество действительных чисел (используемых на практике), функция $Z=f(x_{1,\ldots,} x_n)$, называется *скалярной*. Условия $g_i(x_{1,\ldots,} x_n) \lesseqgtr b_i,\ i=1,m$ предполагающие, что в каждой из n строк может присутствовать одно из рассмотренных ограничений $g_i(x_{1,\ldots,} x_n) \lesseqgtr bi,\ i=1,m$, называются *областью определения* (U) задачи. Функция Z называется *целевой функцией*, которая может представлять собой критерий оптимальности при синтезе какой-либо системы. Таким образом, целевая функция Z достигает экстремального значения в одной или нескольких точек U, которые требуется отыскать.
Обычно в постановке задачи выбора оптимального решения, вид функций Z и g_i, известны, константы b_i определены ,специально оговариваются ограничения, выраженные в требованиях к неотрицательности, либо к целочисленности переменных.

Краткая запись условия задачи математического программирования(М П) имеет вид:

$$X \in U \rightarrow \max(\min)\{Z=f(x)\}$$

Два основных класса задач МП - это задачи линейного и нелинейного программирования.
К задачам *линейного программирования (ЛП)* относятся те, в которых и целевая функция Z, и все функции g_i линейны.
Задачи, в которых присутствуют какие-либо нелинейности, соответственно называются задачами *нелинейного программирования*.

Линейное программирование. Постановка задачи.

Задача ЛП в общей постановке состоит в отыскании переменных $x_{1,\ldots,} x_n$,обращающих в экстремум функцию: $\max \leftarrow Z = \sum_{j=1}^{n} C_j X_j$ при ограничениях:

$$\begin{cases} a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \le b_1 \\ \qquad\qquad \ldots \\ a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \le b_m \end{cases}$$

Обычно, кроме того, на практике добавляют требования не отрицательности$(\forall\ x_j > 0_{ij} = \overline{1,n})$
Пример постановки задачи: Ферма производит откорм скота. Допустим, что имеется 4 вида продуктов (П1,П2,П3,П4). Стоимость единицы каждого продукта (С1,С2,С3,С4). Из этих продуктов требуется составить пищевой рацион, который должен содержать: белков не менее b1 единиц, углеродов не менее b2 единиц, жиров не менее b3 единиц. Для продуктов (П1,П2,П3,П4) содержание белков, жиров и углеродов в единицах на единицу продукта приведено в таблице.

продукты	Элемент		
	белки	углероды	жиры
П1	a_{11}	a_{12}	a_{13}
П2	a_{21}	a_{22}	a_{23}
П3	a_{31}	a_{32}	a_{33}
П4	a_{41}	a_{42}	a_{43}

Где aij некоторые числа и i – продукт j – элемент.
Задача: составить такой пищевой рацион, то есть определить количество продуктов (П1,П2,П3,П4) входящих в рацион, что при минимальной стоимости рациона, при условии выполнения ограничения белков, жиров и углеродов.
$Z = C_1x_1 + C_2x_2 + C_3x_3 + C_4x_4 \rightarrow \min$ здесь мы за $x_1 \ldots x_2$ обозначили количество продуктов которые должны входить в рацион.

$$\begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + a_{13}\, x_3 + a_{14}\, x_4 \geq b_1 \\ a_{12}\, x_1 + a_{22}\, x_2 + a_{32}\, x_3 + 2_{42}\, x_4 \geq b_2 \\ a_{13}\, x_1 + a_{23}\, x_2 + a_{33}\, x_3 + a_{43}\, x_4 \geq b_3 \end{cases}$$

$$\forall x \geq 0$$

Геометрическая интерпретация задачи. Обоснование подхода к решению задачи

Все последующие рассуждения применимы к формулированию задачи ЛП *в стандартной форме*, имеющей следующее представление:

$$Z = \sum_{j=1}^{n} Cjxi \rightarrow \max \qquad (-Z\rightarrow\min)$$

$$\begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1n}\, x_n = b_1 \\ \ldots \\ a_{m1}\, x_1 + a_{m2}\, x_2 + \ldots + a_{mn}\, x_n = b_m \end{cases} \qquad (1)$$

Любую задачу Л П можно свести к подобному виду, в частности, можно изменить знак целевой функции и минимизировать ее. Перейти же от неравенств к равенствам, можно введя дополнительные переменные.
Ответить на вопрос, всегда ли задача имеет решение, а так же сформулировать требования к подходу решения можно на основании *геометрической интерпретации*. Геометрическая интерпретация в постановке и решении задач наглядна при условии n-m=k=2, то есть число уравнений - m на 2 меньше, чем число переменных - n, в этом случае интерпретация может быть проиллюстрирована на плоскости.
Известно, что m линейно-независимых уравнении всегда можно разрешить относительно n базисных переменных, выразив их через k=n-m (в нашем случае 2) свободные переменные. Обозначим свободные переменные – x_1 и x_2, базисные – $x_3 \ldots x_n$, тогда из уравнений (1) можно получить следующую систему:

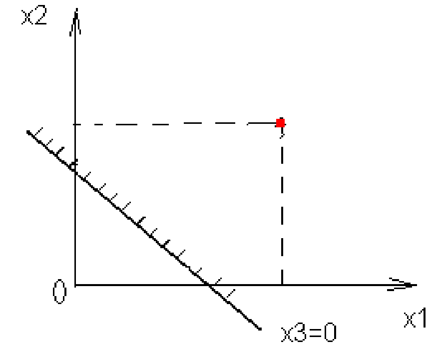
$$\begin{cases} x_3 = a_{31}\, x_1 + a_{32}\, x_2 + B_3 \\ x_4 = a_{41}\, x_1 + a_{42}\, x_2 + B_4 \\ \ldots \\ x_n = a_{n1}\, x_1 + a_{n2}\, x_2 + B_n \end{cases}$$

Теперь будем изображать пару свободных переменных в системе координат.

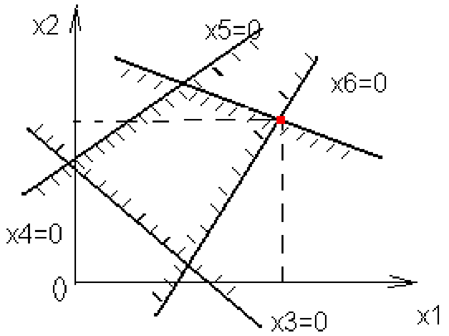
Так как по определению переменные не отрицательны, то первое условие допустимой области решения является первый квадрат системы координат. Теперь уточним область допустимых решений на плоскости $x_1 0 x_2$ с учетом заданных ограничений. Приравняем значение переменной $x_3=0$ и запишем исходное уравнение в следующем виде:

$$0 = x_3 = a_{31}\, x_1 + a_{32}\, x_2 + \beta_3$$

В системе координат $x_1 0 x_2$, данное уравнение, отображается соответствующей прямой (см. рис.). Ее наклон и расположение определяется соответствующими коэффициентами. Данная прямая характеризуется тем, что любая расположенная на ней точка соответствует условию $x_3=0$. Соответственно, в зависимости от коэффициентов, значение x_3 будет больше 0 (выше отсекающей прямой), меньше 0 (ниже отсекающей прямой).



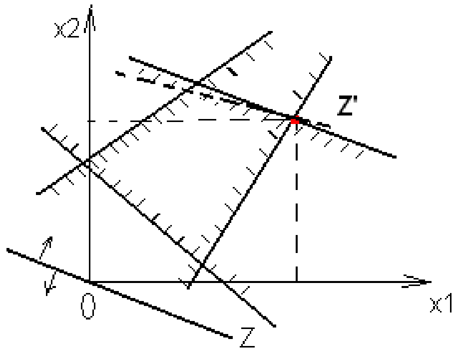
Аналогичным образом можем построить соответствующие прямые для переменных $x_4 \ldots x_n$



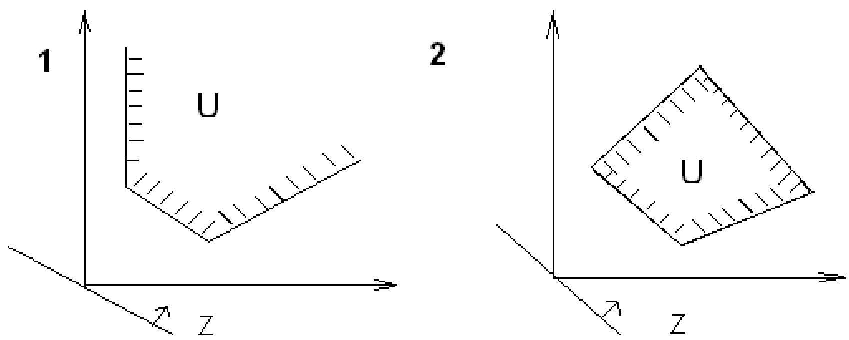
Совокупность данных прямых, при условии неотрицательных переменных, отсекают область допустимых решений (если она не пуста). Проведем подобную операцию с целевой функцией (Z). Выразим Z через свободные переменные и получим:

$$Z = \gamma_1\, x_1 + \gamma_2\, x_2 + \gamma_0$$

Где γ_1 ,γ_2 – коэффициенты, γ_0 – свободный член (который для простоты опустим). Обратим данное уравнение в 0. Тогда, соответственно, для данного уравнения в системе координат $x_1 0 x_2$ может быть построена прямая, проходящая через систему координат. Данная прямая называется *опорной прямой*, для любой точки расположенной на ней, целевая функция равна 0. Обозначим → ← движение, в направлении которого целевая функция возрастает или убывает.



Из данной интерпретации видно, что в общем случае максимальное значение целевой функции достигнет в крайней точке области допустимых решений (в точке пересечения прямых, определяющих область допустимых решений).
Ответим на вопрос, всегда ли задача имеет решение и всегда ли оно единственно. Для этого рассмотрим 2 случая:



1: область не ограничена сверху. Задача не имеет решения, так как всегда можно найти значение лучше предыдущего.
2: ограничение параллельно прямой, отображающей целевую функцию. В этом случае имеется бесконечное множество решений, каждое из которых отображается точкой на соответствующей прямой области определения – это оптимальные решения.
Замечание: важно то, что к оптимальным относятся и краевые точки, задаваемые пересечением прямых области ограничения.

Выводы, основанные на геометрической интерпретации задачи

Рассмотренная геометрическая интерпретация в постановке задачи Л П, позволяет сделать важные выводы относительно подхода к решению задач.

1. решением задачи будет краевая точка, характеризуемая тем, что k=n-m свободных переменных (в нашем случае k=2, x_1,x_2) = 0.
2. как следствие, возможность решения задачи состоит в переборе краевых точек.
3. переход от одной краевой точки к другой состоит в смене свободных переменных.

Очевидно, чтобы пытаться решить задачу, можно простейшим способом задать свободные переменные и решить систему уравнений. Это связано со следующими недостатками:

1. можно получить решение не в области допустимых решений.
2. задача может быть весьма громоздкой, например, n=30 m=10 C_{30}^{10}=30045015.

Вывод: возникает задача направленного перебора краевых вершин, то есть направленного перехода к оптимальному решению. Как следствие возникает задача обоснования критерия оптимальности перехода между краевыми точками.

Симплекс алгоритм (С А) и симплекс метод (С М).

Канонический вид (К В). Возможность простого получения решения задачи

С М может быть применен в том случае, когда задача Л П задана, в так называемом, каноническом виде (каждую задачу можно привести к такому виду).
Канонический вид характеризуется специальным представлением системы ограничения. Система задана в каноническом виде в том случае, если для m-первых переменных входят только в соответствующую строку, причем с коэффициентом равным нулю.

$$\begin{cases} x_1 + a_{1m+1} x_{m+1} + \ldots + a_{1n}x_n = B_1 \\ \qquad \ldots \\ x_m + a_{mm+1} x_{m+1} + \ldots + a_{mn}x_n = B_m \end{cases} \qquad (2)$$

Это связано с тем, что именно при такой постановке задачи строго обоснован алгоритм поиска оптимального решения, реализуемый СА-ом.
К такому виду система всегда сводится, ввиду линейности ограничений, так как всегда можно выразить переменные $x_1 \ldots x_m$ через какие-либо коэффициенты.

$$\begin{matrix} \Psi_1 \\ \\ \Psi_m \end{matrix} \begin{cases} x_1=\Psi_1^*(x_{1+1},\ldots, x_n) \\ \qquad \ldots \\ x_m = \Psi_m * (x_{m+1},\ldots, x_n) \end{cases}$$

В данном представлении, переменные $x_1 \ldots x_n$, называются *свободными*, все остальные *базисными*.
Соответственно при данном представлении всегда можно определить некоторое базисное решение.

$$\begin{aligned} &x_1=B_1 \\ &: \\ &x_m=B_m \\ &x_{n+1} = \ldots = x_n = 0 \end{aligned}$$

Симплекс алгоритм. Процесс приближения к оптимуму (направленный перебор краевых вершин).

Пусть задача представлена в К В, соответственно сразу получаем некое исходное базисное решение.
Возникают следующие вопросы:

- Оптимально ли и единственно ли оно;
- Если не оптимально, а ранее мы показали, что поиск решения состоит в переборе краевых вершин, то возникает вопрос, как осуществить переход к следующему решению.

Вспоминая, что целевая функция выглядит следующим образом:

$$Z=\sum_{j=1}^{n} C_jX_j$$

Просуммируем соответствующие уравнения в системе (2) умножив их на коэффициент C_j.

$$\begin{cases} C_1x_1 + a_{1m+1}x_{m+1} + \ldots + a_{1n}x_n = C_1B_1 \\ \ldots \\ C_mx_m + a_{mm+1}x_{m+1} + \ldots + a_{mn}x_n = C_mB_m \end{cases}$$

Имеем:

$$\sum_{j=1}^{n} C_jX_j + \sum_{j=m+1}^{n} C\square_jX_j = Z_0 + \sum_{j=m+1}^{m} C\square_jX_j \qquad (3)$$

Где $C\square_j$ - коэффициент при свободных переменных и зависит от значения a_{ij} , $B_{i.}$ По своему смыслу Z_0 есть значение целевой функции для исходного базисного решения.

Теорема:

Допустимое базисное решение $x_i = B_i$ при $i = 1\ldots m$, $x_j = 0$ при $j = m+1\ldots m$, является оптимальным, если коэффициенты $C\square_{\square j}$ в выражении (3) не отрицательны.
Доказательство:
Если переменные $C\square \geq 0$, то при любых значениях (не отрицательных) переменных x_j , разность Z и Z_0 неотрицательна и ее минимум есть 0, следовательно, Z не будет меньше Z_0, что говорит о том, что его нельзя уменьшить.
Следствия:

1. Любое решение системы (1), которому соответствует нулевая сумма $(\sum_{j=m+1}^{n} C\square_jX_j)$ при любых $C\square_{\square j} \geq 0$ и $x_j \geq 0$ будет оптимальным, соответственно при существовании $C\square_j < 0$, решение может быть улучшено.
2. Допустимое оптимальное базисное решение окажется единственным, если все $C\square_j > 0$ (строго больше).

Данные следствия являются сформулированными *признаками оптимальности и единственности оптимального решения.*
Теперь рассмотрим, как осуществить переход к новому решению и сформулируем критерий оптимальности выбора направления перехода.
Пусть среди коэффициентов $C\square_j$ есть отрицательные значения, это значит, что давая приращения, соответственным переменным, целевую функцию можно уменьшить, чтобы перейти к другому целевому решению (как следует из интерпретации) необходимо сделать положительной (ввести в базис) какую-то из переменных $x_{n+1} = \ldots = x_n = 0$, обратив в 0 одну из переменных $x_1\ldots x_n$. соответственно в базис должно вводиться та переменная (ей дается положительное приращение), которая имеет в выражении для целевой функции отрицательный коэффициент.
Обозначим через S тот номер j из n+1,…,n величина коэффициента $C\square_j$ для которого отрицательна и максимальна по модулю.

$$C_s=\min[C\square_j <0]$$

Отсюда бросается в глаза очевидный подход к поиску оптимального решения, который и используется в С А.
Соответственно будем давать положительное приращение к переменной X_s . Давая данное приращение, из уравнений (2), получаем:

$$\begin{cases} x_1 = B_1 - a_{1s}\, x_s \\ \ldots \\ xm = Bm - a_{ms}\, x_s \end{cases}$$

Для Z имеем:

$$Z = Z_0 + C□_s X_s \quad (C□_s < 0)$$

Естественно, что выбрав направление движения, целесообразно максимально увеличить переменную X_s, чтобы получить максимальное выражении в значении целевой функции. Однако возрастание переменной X_s ввиду требования к неотрицательной переменной возможно до тех пор, пока какая-то из свободных переменных не обратиться в ноль, то есть, пока не достигнем граничного значения.
Это произойдет если хотя бы один коэффициент с индексом X_s положителен по условию и среди B_i нет отрицательных. Если все a_{is} положительны, то первая в 0 обратиться та базисная переменная, для которой будет выполняться:

$$X_s = \min [B_i / a_{is}]$$

Это и есть формализованное условие перехода к последующей краевой вершине.

Симплекс метод.

Ранее было показано, что СА предполагает направленный перебор краевых вершин. Мы определились с тем, каким образом осуществлять данный перебор, как выбрать направления перебора, как определить, что в итоге получено оптимальное решение (все это ранее было выражено в формальном виде – приведены соответствующие формулы для расчетов).
Однако остается проблема, состоящая в том, как найти то первое допустимое решение (а это должна быть краевая точка), с которой начинать направленный перебор вершин (т.е. когда уже становится возможным применение СА). Для это цели и служит СМ, состоящий в двухэтапном решении задачи. На первом этапе требуется найти какое-либо решение задачи (исходную краевую точку), на втором – осуществить перебор вершин с целью поиска оптимального решения. При этом будем помнить, что краевая точка легко отыскивается, в случае представления задачи в КВ (см. выше).
С М предполагает в общем случае двухэтапного применения С А.
Первый этап начинают с формального введения в каждую строку стандартной системы искусственной неотрицательной переменной для получения К В (напомним, что при представлении задачи в данном виде сразу же можно определить решение задачи).
Исходная:

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ a_{1m}x_1 + \ldots + a_{mn}x_n = b_m \end{cases}$$

К В:

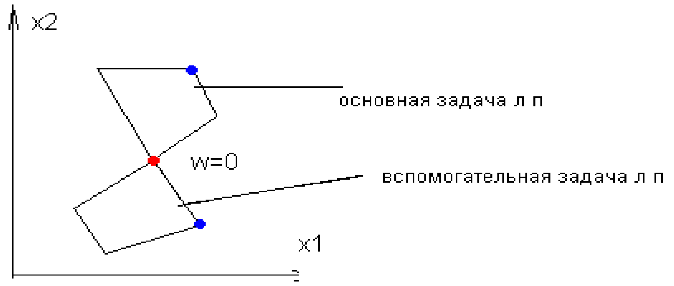
$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n + x_{n+1} = B_1 \\ a_{1m}x_1 + \ldots + a_{mn}x_n + x_{n+m} = B_m \end{cases}$$

Однако этим мы изменили собственно задачу. В каком случае решение новой задачи и исходной задачи будут совпадать?
Если существует решение для первой задачи, то оно допустимо и для второй при условии $x_{n+1} = \ldots = x_{n+m} = 0$ (т.к. именно эти переменные были нами введены, ввиду их неотрицательности, при данном условии все переменные обращаются в 0).
Рассмотрим сумму ($W = x_{n+1} + \ldots + = x_{n+m}$) наименьшее значение которой есть 0 ввиду неотрицательности переменных при $x_{n+1} = \ldots = x_{n+m} = 0$. Таким образом, проверка условия

minW=0 для второй системы ограничений, является условием совпадения решения для первой и второй задач.
С учетом сказанного, первый этап решения С М состоит в поиске какого-либо решения исходной задачи, по средствам решения следующей оптимизационной задачи:

$$minW=\sum_{j=1}^{m} x_{n+j}\,.$$

с учетом ограничений (системы), заданных для второй задачи.
Для данной задачи, приведенной к К В (то есть может быть применен С М), легко находится первое решение. $x_{n+1} = B_1, \ldots, x_{n+m} = B_m$, а $x_1 = \ldots = x_n = 0$. После чего может быть использован С А с целью поиска оптимального решения. Если удается найти оптимальное решение, для которого W=0, это решение является исходным для основной задачи. В противном случае задача не разрешима.



Второй этап состоит в решении основной задачи, в использовании С А уже для решения основной задачи, при этом получим решение вспомогательной задачи рассмотренное как исходное для основной, проверяется на оптимальность при необходимости осуществления последовательного перебора решений (Замечание: здесь уже используется целевая функция Z).

Пример решения задачи ЛП симплекс методом

Пусть имеется некоторый материал в виде стандартных листов, которые надо раскроить таким образом, что бы получить не менее 80 деталей первого типа и не менее 40 второго. Известно 4 способа раскроя листа, для которых результаты раскроя представлены в таблице:

Cn	1	2	3	4
P	3(1)	2(1)	1(1)	0(1)
	1(2)	6(2)	9(2)	13(2)

Требуется получить необходимое число деталей при минимальном расходе листов материала. Обозначим через x_j количество листов раскраиваемых j-ым столбцом. Целевая функция в данном случае:

$$Z = x_1 + x_2 + x_3 + x_4$$

При этом должны выполняться следующие условия:

$$\begin{cases} 3x_1 + 2x_2 + x_3 \geq 80 \\ x_1 + 6x_2 + 2x_3 + 13x_4 \geq 40 \end{cases}$$

$$\forall\ x_j \geq 0,\ j=\overline{1,n}$$

Для применения СМ, прежде всего, задачу требуется привести к стандартному виду.

$$\begin{cases} 3\ x_1 + 2\ x_2 + x_3 - x_5 = 80 \\ x_1 + 6x_2 + 9x_3 + 13x_4 - x_6 = 40 \end{cases} \qquad (4)$$

Таким образом, так как ограничения были $\geq$, из каждой строки нужно вычесть соответственную переменную. Очевидное базисное решение:

$$- x_5 = 80$$
$$- x_6 = 40$$
$$x_1 = \ldots = x_4 = 0$$

недопустимо ввиду не отрицательности переменной. Так как данное решение не допустимо, задачу необходимо решать в 2 этапа.

<u>1 этап:</u>

Решение вспомогательной задачи: приведем исходную задачу к К В:

$$\begin{cases} 3\ x_1 + 2\ x_2 + x_3 - x_5 + x_7 = 80 \\ x_1 + 6x_2 + 9x_3 + 13x_4 - x_6 + x_8 = 40 \end{cases}$$

Для этого введем дополнительные переменные x_7 , x_8 . соответственно формулируем и решаем данную задачу с целевой функцией, которая должна быть минимизирована, помня, что ее решение будет и решением основной, при условии W=0.
Здесь легко отыскиваются исходные значения:

$$x_7 = 80,\ x_8 = 40,\ x_1 = \ldots = x_6 = 0,\ W = 120$$

Применим С А и прежде всего проверим полученное решение на оптимальность.
Рассмотрим целевую функцию $W = x_7 + x_8$, в которой переменные x_7, x_8 выразим через остальные $W = x_7 + x_8 = 80 - 3\ x_1 - 2\ x_2 - x_3 + x_5 + 40 - x_1 - 6x_2 - 9x_3 - 13x_4 + x_6 = 120 - 4\ x_1 - 8\ x_2 - 10\ x_3\ -13\ x_4 + x_5 + x_6$
Присутствие отрицательных коэффициентов указывает на не оптимальность полученного решения, так как максимальное по модулю отрицательное значение коэффициента имеет переменная x_4 , что в соответствии с рассмотренным ранее критерием именно ей будем давать положительное приращение ($x_3 = x_4$).
Теперь рассмотрим, до какой величины можно увеличить x_4 , и какая при этом переменная обнулится.

$$\begin{cases} x_7 = 80 \\ x_8 = 40 - 13\ x_4 \end{cases}$$

Вводя положительную переменную x_4 из исходной системы, для нашего решения имеем: таким образом видим, что x_4 можно увеличить до значения 40/13, при этом x_8 обратится в 0 и получим новое решение:

$x_1 = x_2 = x_3 = 0,\quad x_4 = 40/13,\quad x_5 = x_6 = 0,\quad x_7 = 80,\quad x_8 = 0,\quad W = 80+40/13$

Проверим полученное решение на оптимальность, для этого выразим целевую функцию. Так как $x_8 = 0$ выражаем только x_7 .

$$W = x_7 + x_8 = -3\ x_1 - 2\ x_2 - x_3 + x_5 - x_8$$

Видим, что решение не оптимально, и так как максимально отрицательный коэффициент имеет переменная x_1 , то именно ей надо дать положительное приращение. Определим, до какого значения можно увеличить x_1 , и какая при этом переменная обратится в 0 (x_4 и x_7 – не нулевое значение).

$$x_7 = 80 - 3\ x_1$$
$$x_4 = (40 – x_1)/13$$

x_1 может расти до 80/3 при этом первой обращается в 0 x_7 . получаем:

$x_1 = 80/3$
$x_4 = 40/39$
$x_2 = x_3 = x_5 = x_6 = x_7 = x_8 = 0$ (оптимальное решение)
$W = x_7 = x_8 = 0$

Так как целевая функция не содержит отрицательных коэффициентов и ее минимум есть 0, то найденное оптимальное решение может рассматриваться как исходное решение для выполнения 2 этапа.
2 этап:

$$\begin{cases} 3\ x_1 + 2\ x_2 + x_3 - x_5 = 80 \\ x_1 + 6x_2 + 9x_3 + 13x_4 – x_6 = 40 \end{cases}$$

исходное решение:
$x_1 = 80/3$
$x_4 = 40/39$
$x_2 = x_3 = x_5 = x_6 = x_7 = x_8 = 0$
$W = x_7 = x_8 = 0$

Имея допустимое базисное решение, полученное на первом этапе решения, начинаем решение основной задачи.

<u>2 этап:</u>

Прежде всего, проверим оптимальность полученного решения, для этого выразим x_1 и x_4.
Получаем из (4):

$x_1 = (80 - 2x_2 - x_3 + x_5) / 3$
$x_4 = (40 - x_1 - 6x_2 - 9x_3 + x_6) / 13$

Целевая функция: $Z= x_1 + x_2 + x_3 + x_4 = (1080 - 3\ x_2 + 12x_1 + 3\ x_6) / 39$
Вывод:

- Полученное решение не оптимально.
- Следует дать приращение переменной x_2 . Для определения того, на сколько увеличивать x_2, выразим наши нулевые переменные x_1 и x_4 через x_2 получим:

$$x_1=(80-2\ x_2)/3$$
$$x_4=(40-16\ x_2)/39$$

x_1 выражаем через x_2. Отсюда получаем, что x_2 может увеличиваться до 2,5, при этом x_4 превращается в 0.

$x_1 = 25$, $x_2 = 2{,}5$, $x_3 = x_6 = 0$ целевая функция: $Z=27{,}5$

Проверим данное решение на оптимальность, для этого выразим целевую функцию через переменные: $Z= x_1 + x_2 + x_3 + x_4 = (440+2x_3+3x_4+x_5+5x_6)/16$

Вывод:
Данное решение задачи является оптимальным и единственным. Таким образом, оптимальный способ раскроя листов (см задачу) состоит в следующем: 25 кроить 1-ым способом, 2,5 – вторым, в результате потребуется 27,5 листов.

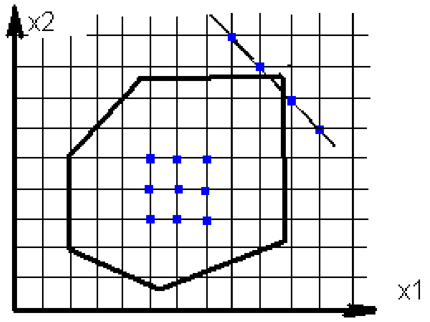
Недостатки С М (метода решения, основанного на геометрической интерпретации задачи ЛП):

- ***Собственно критерий оптимальности не призван оптимизировать вычислительную процедуру, так как сложность вычислений определяется числом перебираемых вершин, в то время, как критерий оптимальности не минимизирует их числа при поиске оптимального решения.***
- ***Сложность использования вычислительных средств для решения задачи. Алгоритмы, ориентированные на применение вычислительных средств, должны характеризоваться по возможности простейшими операциями (+,- и т.д.). Здесь же при обработке каждого решения, необходимо решать достаточно сложные системы уравнения, что сложно формируется на ЭВМ. Таким образом, можем сделать вывод, что для решения задач большой размерности, где необходимо применение ЭВМ, необходимы иные подходы, следовательно, другая интерпретация задачи и критерий оптимальности.***
- ***Особенно сложен С М для решения целочисленных задач.***

Целочисленное программирование.

Метод Гомори (М Г).

Задачи сформулированные в терминах Л П и содержащие требования, все или некоторые переменные, целые числа называются *задачами целочисленного программирования*. Рассмотрим пример интерпретации задачи:



Пусть к x_1 выставлены требования целочисленности. Решения задачи в этом случае будут располагаться на соответствующих не пересекающихся линиях, характеризующие целочисленность переменных. Если обе переменные целочисленные, то решение задачи определяется узлами соответствующей решетки на плоскости.
М Г состоит из выполнения следующих шагов:

1. Решается задача классическим С М и полученное решение проверяется на целочисленность.
2. Если полученное решение не целочисленное, вводят по определенным правилам в систему ограничений отсекающее полученное решение и задача решается вновь.
3. Процедура повторяется, пока не найдено целочисленное решение.

Вывод:
Условие целочисленности принципиально усложняет решение задачи, так как в общем случае требует многократного применения С М, с увеличением области ограничения (с введением дополнительных условий).

Методы оптимизации, ориентированные на применение ЭВМ. Общий подход к построению

Сложность задач М П вызывает необходимость разработки методов оптимизации, базирующихся на вычислительных и логических операциях. Общим для этих методов является необходимость доказательства целесообразности и достаточности. Целесообразность обусловливается интерпретацией или физическим смыслом метода решения, достаточность доказательством сходимости алгоритма.
Каждый алгоритм предполагает некоторое правило поиска решений и обусловливает переход к последующему решению. Отношения, связывающие очередные решения при использовании алгоритма называется *алгоритмическим отображением*. *Алгоритм*, построенный на основе отображения, *сходится*, если он позволяет либо непосредственно вычислить оптимальное (экстремальное) значение, либо представляет его в виде некоторой предельной точки задаваемой последовательностью приближений.
В общем случае следует говорить о сходимости алгоритма в смысле существования подобной предельной точки. Основной группой методов оптимизации, ориентированных на использование ЭВМ, являются так называемые *методы возможных направлений*. Общая идея методов возможных направлений состоит в следующем: от изначально допустимой по условию задачи точки x_1 осуществляется переход к новой допустимой точке, в которой значение целевой функции лучше, чем в предыдущей. Процесс продолжается до тех пор, пока остается возможность улучшать значение целевой функции, либо задача не может быть решена с заданной точностью. При этом на каждом шаге, для перехода в последующую точку, требуется определить направления перехода и «длины шага». Таким образом, следующий шаг определяется по формуле:

$$x_{k+1} = x_k + \alpha_k r_k \qquad (5)$$

где α_k – величина k-ого шага
r_k – единичный вектор, определяющий направление перемещения.

Данное равенство является общим для рассматриваемой группы методов оптимизации, интерпретация же подхода к ее решению призвана задать правило определения параметров r_k и α_k.

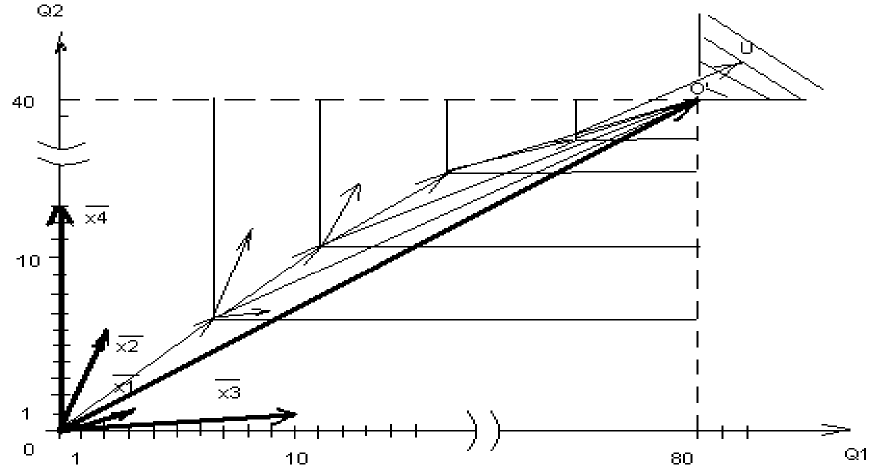
Векторный метод решения задачи Л П.

Рассмотрим данный метод, как метод, относящийся к группе методов возможных направлений, оценим его преимущества и недостатки по сравнению с традиционным С М. Проиллюстрируем интерпретацию данного метода решенным ранее примером:

Min $\{Z=\Sigma x_j\}$

$$\begin{cases} 3x_1 + 2x_2 + 10x_3 \ge 80 \\ x_1 + 6x_2 + x_3 + 13x_4 \ge 40 \\ \forall\, x_j \ge 0,\ j=1,4 \end{cases}$$

В данном случае интерпретацию на плоскости можно дать, если число ограничений в системе не превышают двух. Отложим в системе координат шкалы ограничений.

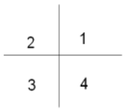


Данные ограничения позволяют задать допустимую область решений U. Естественно так как требуется достигнуть данного ограничения с минимальными потерями (нас интересует оптимальное решение), что обеспечивается кратчайшим путем между точками O и O', следует провести вектор OO', задающий кратчайший путь к области допустимых решений. Каждое возможное решение, задаваемое системой ограничений, так же можно представить собственным вектором на данной плоскости. В частности, если принять $x_1 = 1$ (x=3, y=1, см. исходную задачу), ему будет соответствовать вектор x_1. Соответственно получаем все остальные вектора. Очевидно, что для заданных ограничений лучшим был бы вектор или решение, которое позволяет продвигаться со скоростью 2/1 или совпадающим по направлению с идеальным вектором OO'. Однако для заданной задачи подобного вектора нет, $\Rightarrow$ возникает задача выбора оптимального решения (выбора параметра r_k в уравнении (5)).
Так как вектор в общем случае характеризуется 2-мя параметрами (длинной и направлением), в данном случае величиной угла между им и идеальным вектором OO', что с определенными оговорками можно свести к одному критерию выбора лучшего вектора по длине его проекции на идеальный вектор OO'. Итак, по величине длины проекции можно выбрать лучший вектор на данном шаге (в данном случае в качестве α_k предполагаем 1 – один отложенный вектор). Отложение выбранного лучшего вектора на данном шаге (что означает присвоение соответствующей переменной x_i значения 1) приближает нас к допустимой области решений $\Rightarrow$ для перехода к следующему шагу оптимизации, следует перенести начало системы координат в направлении и на величину выбранного лучшего вектора. Соответственно строится и новый идеальный вектор, учитывающий достижение заданных ограничений на предыдущих этапах. В новой системе опять откладываются исходные векторы, и вновь выбирается лучший вектор. Данная процедура выполняется до тех пор, пока последний лучший из отложенных векторов не пересечет допустимую область решений. Совокупность отложенных векторов на всех этапах и определяет оптимальный план или оптимальное решение.
Замечание: видим, что в данном случае мы получаем оптимальное решение целочисленной задачи.

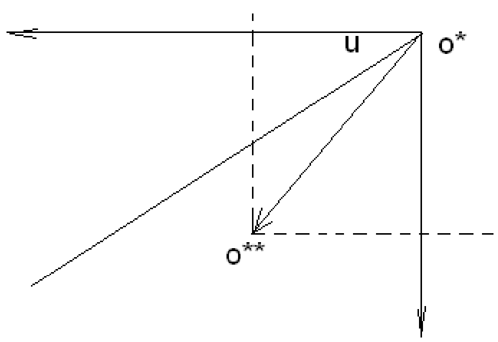
Подходы к точному решению задачи.

1. Метод последовательных приближений.

Идея метода состоит в том, что на первом шаге задачи решаются целочисленно (откладываются полные вектора). При пересечении последним отложенном вектором допустимой области решения – U, начинается процесс приближения к оптимальному точному решению. Каждый этап такого приближения состоит в следующем: изменяется исходная система координат (с 1-ого квадранта на 3-ий, со 2-ого на 4-ый).



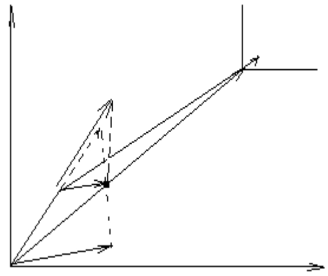
Из точки пересечения последнего отложенного вектора допустимой области решений – U, формируем идеальный вектор относительно которого теперь выбираем лучший вектор. Векторы вариантов изменяют направление на противоположные и мы начинаем рассматривать некоторые их части (длины векторов умножаются на коэффициенты меньше 1). Процедура повторяется до тех пор, пока последний отложенный вектор (часть вектора), не достигнет точки O**



2. Метод точного решения.

Точное решение задачи: данный альтернативный подход предполагает отложение на каждом шаге не целиком, а лишь часть векторов. Данная часть выбирается из следующих соображений. Предположив, что после отложенного лучшего вектора, следующим лучшим становится другой вектор, всегда можно определить ту часть первого вектора, которая задает лучшее направление продвижения в сторону исходного вектора. Именно эта часть, а не целиком вектор. Искомая часть вектора может быть определена следующим способом: если из конца искомой части лучшего вектора отложить идеальный вектор, то отложенный из него 1-ый и 2-ой лучшие вектора будут иметь равную длину проекции на вновь отложенный идеальный вектор.

Данный метод уже трудно отнести к методам, ориентированным на ЭВМ, так как параметр α_k определяется по средствам решения уравнений.

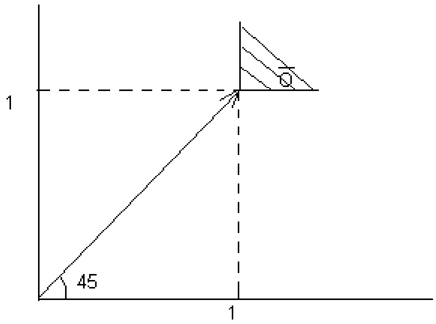


Формализация решения задачи векторным методом

Как отмечали, критерий выбора решения на каждом этапе определяется двумя параметрами – длиной вектора решения и величиной угла вектора решения с идеальным вектором (n=0). Соответственно свести задачу к однокритериальной, можно рассматривая длину проекции вектора решения на идеальный $|x_{no}| = x_n \cos\alpha_{no}$

Если обозначить через $x_{n1} = \overline{1,n}$ параметры n-ого варианта, то длина вектора равна: $x_n = [\sum_{i=1}^{m} x^2_{ni}]^{1/2}$

Ортонормируем систему следующим образом. Примем $\overline{x_{ni}} = x_{ni} / x_{oi}$, $i=\overline{1,n}$



$\overline{X_{oi}} = 1$
$i=\overline{1,n}$

Получим длину проекции в ортонормированном базисе. $x_{no} = (1/\sqrt{m}) \sum_{i=1}^{m} \overline{x_{ni}}$

Замечание. Длина вектора и угол между векторами определяются по следующим формулам:

$$x_n = \left[\sum_{m=1}^{M} q_{nm} \right]^{\frac{1}{2}}$$

$$\alpha = \arccos\left[\sum_{m=1}^{M} q_{0m} q_{nm} \left(\sum_{n=1}^{M} q^2{}_{0m} \sum_{m=1}^{M} q^2{}_{nm} \right)^{\frac{1}{2}} \right]$$

Данное выражение после соответствующих подстановок, в предположении, что нормированное значение идеального вектора равно 1:

$$q_{0m} = 1 \quad \Rightarrow \quad x_{n0} = \frac{1}{\sqrt{M}} \cdot \sum_{m=1}^{M} \overline{q}_{nm}$$

Так как значение m фиксировано для всех вариантов решения, это значение можно опустить, таким образом, в качестве критерия выбора оптимального вектора, на каждом шаге, можно рассматривать параметр, определяемый следующим выражением:

$$\begin{cases} x_{no} = \sum\limits_{i=1}^{m} \overline{x}_{ni} \\ \overline{x}_{ni} = x_{ni} / x_{oi} \end{cases} \qquad (6)$$

Видим, что эти формулы просты, что позволяет эффективно использовать данный метод решения задачи ЛП с использованием вычислительных средств.
Таким образом, методика выбора оптимального решения сводится к следующему:

1. на каждом шаге используем критерий (6) , выбирается лучшее значение исходя из

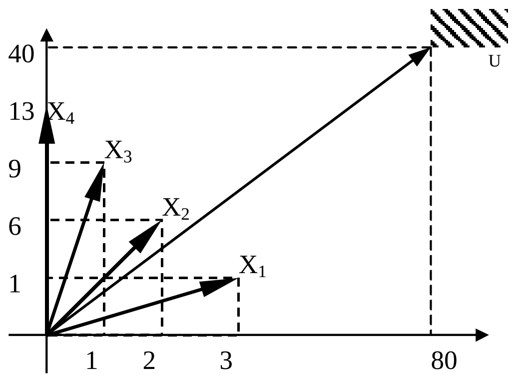
   $x_{no} = \sum\limits_{i=1}^{m} \overline{x}_{ni} \to \max$

2. с учетом выбранного вектора (что соответствует его откладыванию в системе координат) изменяем систему ограничений (вычитаем из ограничений соответствующие параметры векторов).
3. проверяем все полученные ограничения на условие ≤ 0. если для всех ограничений оно выполняется, задача решается. Если условие не выполняется возвращаемся к шагу 1, с перенормированием значений параметров.

Пример: вернемся к задаче раскроя материала, рассмотренной ранее.

$z = \sum x \to \min \quad Q = 80 \quad Q = 40$

Способ	1	2	3	4
Раскрой первого типа	3	2	1	0
Раскрой второго типа	1	6	9	13



$3x_1 + 2x_2 + x_3 \geq 80$
$x_1 + 6x_2 + 9x_3 + 13x_4 \geq 40$

Для иллюстрации выполним несколько шагов решения задачи.

Шаг 1. Считаем значения критерия для различных вариантов.

$1: \varphi_{A_n} = 1 \cdot 1 + \frac{1}{13} \cdot \frac{1}{2} \approx 1{,}04$

$2: \varphi_{A_n} = \frac{2}{3} \cdot 1 + \frac{6}{13} \cdot \frac{1}{2} \approx 0{,}9$

$3: \varphi_{A_n} = \frac{1}{3} \cdot 1 + \frac{9}{13} \cdot \frac{1}{2} \approx 0{,}7$

$4: \varphi_{A_n} = 0 \cdot 1 + 1 \cdot \frac{1}{2} \approx 0{,}5$

Сравниваем – выбираем лучший вариант – это вариант1 (присваиваем $x_1 = 1$).

Формируем новые ограничения: $\begin{cases} Q_1 = 80 - 3 = 77 \\ Q_2 = 40 - 1 = 39 \end{cases}$. Задача не решена, следовательно, требуется следующий шаг решения (не все ограничения нулевые). Следующий шаг отличается от предыдущего тем, что в аддитивном критерии изменяется.

Шаг 22. Рассчитываем значение критерия $\begin{array}{l} 1{:}\varphi_{A_n} = 1\cdot 1 + \frac{1}{13}\cdot\frac{39}{77} = \ldots \\ 2{:}\varphi_{A_n} = \frac{2}{3}\cdot 1 + \frac{6}{13}\cdot\frac{39}{77} = \ldots \\ 3{:}\varphi_{A_n} = \frac{1}{3}\cdot 1 + \frac{9}{13}\cdot\frac{39}{77} = \ldots \\ 4{:}\varphi_{A_n} = 0\cdot 1 + 1\cdot\frac{39}{77} = \ldots \end{array}$. И так далее (выбираем лучшее значение по данному критерию, фиксируем отложенный x, пересчитываем ограничения, проверяем, решена ли задача полностью или требуется следующий шаг решения). И т.д.

Нормирование параметров в общем случае

В общем случае нормирование параметров должно учитывать:

1. Величину вектора относительно векторов других сравниваемых вариантов или на сколько велик шаг в сторону ограничения будет при откладывании данного вектора.
2. Отношение ограничений между собой или в какой мере следует продвигаться при откладывании векторов к конкретному ограничению.
3. Учет отношения влияния выбранного решения по отношению к остальным на целевую функцию.

Нормированное значение параметра задается объективным способом нормирования по следующим условиям: $\bar{q}_{n,m} = \dfrac{q_{n,m}}{q_{0,m}}$; $\quad q_{0,m} = \sup\left\{q_{n,m}, n = \overline{1,N}\right\}$. SUP – верхняя точечная граница (максимум). Учет отношения ограничений осуществляется внесением в весового элемента, определяемого по формуле: $\beta_m = \dfrac{Q_m}{Q_m^{\max}}$; $\quad Q_m^{\max} = \sup\left\{Q_m, m = \overline{1,M}\right\}$.

Таким образом, выбор лучшего решения на каждом этапе выполняется по формуле: $\varphi_{A_n} = \sum\limits_{m=1}^{M} \dfrac{q_{n,m}}{q_{0,m}} \cdot \dfrac{Q_m}{Q_m^{\max}}$. Для того чтобы учесть, на сколько выбранное решение влияет на целевую функцию (при условии, что она минимизируется) в критерий для варианта n следует ввести коэффициент γ_n.

Тогда, получим:

$$\varphi_{A_n} = \gamma_n \sum_{m=1}^{M} \frac{q_{n,m}}{q_{0,m}} \cdot \frac{Q_m}{Q_m^{\max}}, \text{ где } \gamma_n = \frac{c_n^{\min}}{c_n}; \quad c_n^{\min} = \inf\left\{c_n, n = \overline{1,N}\right\}.$$

Эта формула является основной расчетной формулой при выборе оптимального плана на каждом шаге при постановке задачи в общем виде (необходимо учитывать все коэффициенты).

Достоинства векторного метода (метода решения, основанного на векторной интерпретации задачи ЛП):

Кроме очевидной наглядности данная интерпретация позволяет:

- ***В отличие от СМ векторный метод позволяет минимизировать число вычислений, так как критерий оптимальности позволяет на каждом шаге наилучшим образом (с точки зрения количества вычислений) продвигаться в сторону области допустимых решений.***

- ***Минимальна вычислительная сложность на каждом этапе выбора плана (простейшие арифметические и логические операции).***

Все это обусловливает эффективность применения метода для расчета оптимизационных задач на вычислительных машинах.

Приоритетная задача, линейного программирования

Приоритетная задача состоит в том, что в рамках заданных ограничений дополнительного оговаривается порядок выполнения данных ограничений. Например, возвращаясь к нашему примеру раскроя материала, где заданы ограничения - не менее 80 деталей первого типа и не менее 40, может быть дополнительно введено, например, следующее ограничение - в первую очередь, необходимо произвести не мерее 30 деталей первого типа и не менее 20-второго, во вторую очередь……. – и т.д., т.е. вводится некоторый приоритет на очередности решаемых задач. При такой постановке задачи уже действительно можно говорить об оптимальном планировании, т.к. учитывается собственно последовательность выполнения работ.

При использовании классических методов решении задач линейного программирования (СМ), не позволяющих учитывать динамику плана, данная задача разбивается на соответствующую совокупность не взаимосвязанных задач ЛП – число задач определяется числом приоритетов.

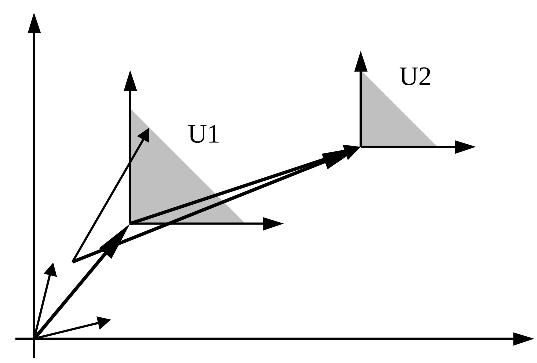
Векторный метод решения задач ЛП позволяет строить оптимальные планы с учетом динамики выполнения работ. Рассмотрим, как решается приоритетная задача линейного программирования векторным методом, и, в первую очередь, как она интерпретируется.

В общем случае приоритетная задача линейного программирования в своей формальной постановке выглядит следующим образом: пусть задано I приоритетов на заданную область ограничений. Соответственно имеем следующую систему ограничений: $I, Q_{mi}, m = \overline{1, M}, i = \overline{1, I}$. При этом задается, что в I-ю очередь требуется выполнить Q_m ограничений. Постановка задачи имеет вид: найти $x_{n,i}$, такие, что $x_{n,i}, n = \overline{1, N}, i = \overline{1, I}$.

Обращается в минимум целевая функция: $\sum_{n_i=1}^{N_i} x_n \to \min, i = \overline{1, I}$ При условии,

что: $\begin{matrix} q_{11}x_{1i} + q_{12}x_{2i} + \ldots + q_{1N}x_{Ni} \geq Q_{1i} \\ q_{M1}x_{1i} + q_{M2}x_{2i} + \ldots + q_{MN}x_{Ni} \geq Q_{Mi} \end{matrix}$. При чем: $\sum_{i=1}^{I} Q_{mi} = Q_m, m = \overline{1, M}$.

Векторная интерпретация задачи состоит в следующем (см. рис.). На плоскости

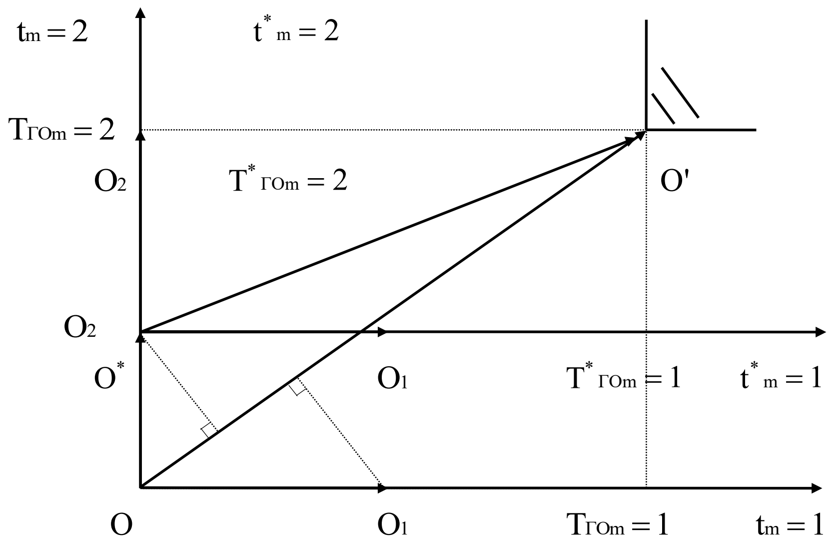


появляется система допустимых областей решений. Допустимая область с меньшим порядковым номером более приоритетна, данные ограничения должны выполняться в первую очередь. В каждой допустимой области должен быть отложен идеальный вектор из начала координат в первую область и из вершины предыдущей области в начало координат во второй области. Далее задача может решаться так же как и в случае без приоритетной постановке, путем выбора лучшего вектора на каждом этапе. Однако существуют два правила преодоления области допустимых решений.

Таким образом, данный подход позволяет использовать возможности векторной интерпретации (возможность поиска не только лучших решений в плане, но и их очередность) для решения приоритетных задач, отличающимися своей постановкой - заданием очередности выполнения ограничений.

Задача синтеза расписаний. Векторная интерпретация

Большинство задач прикладной области вычислительных систем сетей, являются линейными $\Rightarrow$ задачи векторной интерпретации, имеют достаточно большую общность. В частности, всегда, когда может быть определено некое желаемое или оптимальное решение, могут быть определены исходные способы решения, при необходимости задана область оптимальных решений, применим данный подход. Рассмотрим его применительно к задачам синтеза расписания реального времени.
Рассматриваемая детерминированная задача теории расписаний состоит в определении очередности передачи прав на занятие ресурса, минимизирующей его производительность при выполнении ограничений реального времени.
Рассмотрим интерпретацию на рисунке.



По осям координат откладываются временные характеристики решения задачи m-го типа (в частности, t_m) и ограничения реального времени $T_{ГОm}$, которые ограничивают некоторую допустимую область решений (однако, в отличие от предшествующего случая (задача линейного программирования), допустимая область решений здесь уже иная, задается условиями $t_m \le T_{ГОm}, m = \overline{1, M}$). Особенностью является и то, что исходно задано M способов решения задачи, при каждом весь квант времени (с учетом затрат времени на выбор заявки из очереди) предоставляется заявке одной из M очередей. Соответственно имеем M векторов равной длины, совпадающих по направлению с осями системы координат - векторы $< OO_1 >$ и $< OO_2 >$ (см. рис.). В этом случае уже имеет смысл говорить не о максимальной длине проекции на идеальный вектор $< OO' >$, а о минимальной, что реализуется при откладывании векторов по осям координат, так как целесообразно обслужить максимальное число заявок за ограниченное время, или о достижимости ограничений $T_{ГОm}$ по M осям.

Естественно, что при синтезе расписания, первую очередь, имеет смысл откладывать вектор в том m-м направлении, которое может быть получено из условия $T_{ГОm} = \inf\{T_{ГОm}, m = \overline{1,M}\}$, т.е. выбираем самое «жесткое» ограничение, которое должно быть выполнено за минимальное время. На рис. это ограничение $T_{ГОm=1}$. Предоставим ресурс выбранной заявке на фиксированное время занятия, задаваемое продолжительностью временного кванта, что на рис. соответствует отложению вектора $< OO_1 >$. Далее требуется сформировать новую систему координат. Особенностью ее получения будет перенос системы координат в направлении и на величину векторов $< OO_i, i = \overline{2,M} >$, что на рис. соответствует переносу на вектор $< OO_2 >$.
Поясним физический смысл получения новой системы координат. После обслуживания заявки m=1 ее место занимает следующая заявка из этой очереди, для которой ограничение на время обслуживания совпадает с исходным. Ни одна из заявок других очередей за это время не обслуживается или после окончания кванта ограничение на время обслуживания этих заявок уменьшится ровно на продолжительность кванта. Многоэтапная процедура оптимизации выполняется до тех пор, пока заявки всех очередей не будут обслужены за ограниченное время, и результатом решения задачи будет оптимальное расписание обслуживания при заданной длине кванта, которая, собственно, и задает производительность ресурса. Если заявки всех очередей невозможно обслужить с учетом заданных временных ограничений, необходимо повышать производительность ресурса, что приведет к уменьшению длины кванта, задающего продолжительность обслуживания одной заявки. Затем процедура синтеза расписания должна быть выполнена вновь. Если заявки всех очередей можно обслужить за время, меньшее задаваемого ограничениями, можно увеличить продолжительность кванта с повторным выполнением процедуры для оптимизации производительности ресурса.
Утверждение. Продолжительность временного кванта обслуживания заявки в системе T_K задается условием $T_K \le \inf\{T_{ГОm}, m = \overline{1,M}\} / M$.

Доказательство. Чтобы иметь возможность обслужить M заявок с совпадающими $T_{ГОm}$ (худший для системы случай - бесприоритетная дисциплина обслуживания), необходимо и достаточно, чтобы за время $T_{ГОm}$ каждая заявка единожды получила доступ к ресурсу, что возможно при условии выбора продолжительности кванта, равной $T_{ГОm} / M$.

Таким образом, использование векторного метода синтеза расписаний состоит в следующем - выбирая очередность представленного кванта - синтезируем дисциплину обслуживания; выбирая же величину кванта – определяем оптимальную производительность ресурса.

Многокритериальная оптимизация. Векторная интерпретация задачи.

На практике практически отсутствуют системы, характеризуемые одним единственным параметром. Поэтому для сравнительного анализа необходимо учитывать одновременно несколько параметров, характеризуемых разнородностью (производительность, надежность, стоимость, габариты и т.д. и т.п.) Как же произвести подобное сравнение? С этой целью используются методы многокритериальной оптимизации.
Данный класс задач предполагает сравнение вариантов решения одновременно по нескольким параметрам, где каждый n–ый вариант характеризуется m-ым числом параметров:

n, n=1,N
m, m=1,M
(q_{nm}, n=1,N, m=1,M)

Для сравнения вариантов между собой необходимо сведение задачи к однокритериальной. Существует два основных подхода, различающихся типом «свертки» параметров. Это мультипликативная и аддитивная свертка.

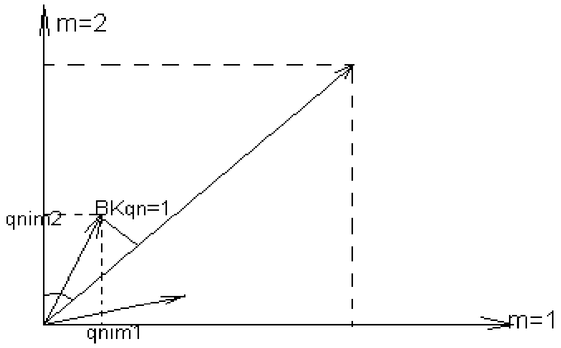
$$Фn = \prod_{m=1}^{M} q_{nm} \text{ – мультипликативная}$$

$$Yn = \sum_{n=1}^{N} \overline{q_{nm}} \text{ - аддитивная}$$

Первый подход предполагает простое перемножение параметров между собою. Не говоря об отсутствии обоснованного физического смысла (производительность умножается на надежность) данный подход характеризуется следующим огромным недостатком – существенная зависимость агрегатирующего критерия оптимальности от отдельных локальных качеств (если умножить на 0, то результатом будет 0).
В связи с этим на практике широкое применение нашла аддитивная свертка. Для ее применения необходимо решить два вопроса:

1. Как учесть разный физический смысл (природу) различных параметров.
2. Как учесть различную относительную важность разнородных параметров.

Дадим векторную интерпретацию задачи. Рассмотрим данную задачу на плоскости:



Каждый вариант решения на плоскости образует вектор качества. В рамках данной интерпретации получаем, что вариант решения характеризуется длиной вектора качества и величиной угла с неким идеальным вектором, или длиной проекции на идеальный вектор:

$$\left.\begin{array}{l} |Xn| \to max \\ \\ \alpha no \to min \end{array}\right| \to Xno = Xn \cos\alpha no$$

Т.е., в предположении, что задан некоторый идеальный вектор качества (q_n=0), лучший вариант можно определить по длине вектора и углу вектора с идеальным. Или сводя задачу к однокритериальной по длине проекции вектора качества на идеальный вектор. Ранее мы показывали, что в ортонормированном базисе данная длина определяется следующей функцией:

$$Xm = (1/\sqrt{Mm}) * \sum_{m=1}^{M} \overline{q}_{nm} = Yn = \sum_{n=1}^{N} \overline{q}_{nm}$$

а это ни что иное, как аддитивная свертка.
Таким образом, аддитивный критерий оптимальности может интерпретироваться как приведенная (умноженная на коэффициент) длина проекции вектора качества варианта на идеальный вектор качества.

Возникают 2 вопроса:
1. Как нормировать параметры.
2. Как учитывать относительную важность.
Какова интерпретация решения этих задач.

Объективный способ нормирования.
Суть его состоит в оценке достигаемости сравниваемыми вариантами лучших значений параметров, определяемых на сравниваемом множестве вариантов.
Реализуется нормирование следующим образом:
$\overline{q}_{nm} = q_{nm} / q_{nmax}$, если лучшее значение качества определяется, как max (например, производительность)
$\overline{q}_{nm} = q_{nmin} / q_{nm}$, если лучшее значение качества определяется, как min (например, стоимость)
Подобное нормирование «наполняет» аддитивную свертку определенным физическим смыслом – сложению подвергаются не собственно разнородные параметры, а меры достижения каждым вариантом некоего лучшего значения, определяемого на сравниваемом множестве сравниваемых вариантов.
Достоинства:
- Просто формализуется решение в части выбора лучших значений параметров.

Недостатки:
- Гипотетичность идеального вектора качества по существу не достижима.
- Жесткие требования к анализу сравниваемого множества решений (например, введите в сравниваемое множество вариантов компьютеров счеты – и считайте, что на сравниваемом множестве вариантов вы отказались от учета таких параметров, как цена и надежность).
- Актуальной, по прежнему, остается задача учета относительной важности параметра.

Данная задача решается вводом весовых коэффициентов.

$$\overline{q}_{nm} = \sum_{m=1}^{M} \alpha m * \overline{q}_{nm}$$

Заметим, что подобная эвристика сводит на нет сам метод. Здесь присутствует подмена одной неопределенности другой, и на вопрос, что проще, сравнить между собою варианты, либо численно оценить отношения важности параметров между собой (заданием весовых коэффициентов) даже ответить невозможно.
Однако вернемся к интерпретации. Что с этой позиции означает задание весового коэффициента в ортонормированном базисе - поворот в пространстве идеального вектора (это нам далее пригодится). А идеальный вектор в данном случае определяется лучшими значения параметров на сравниваемом множестве вариантов (в ортонормированном базисе эти значения равны «1»).

Субъективное нормирование.
При проектировании любой системы требуется достижение некоторых качеств (в противном случае речь не может идти о проектировании, а тогда что и зачем сравнивать), совокупность данных требуемых качеств и определяет идеальный вектор, то есть параметры q_{nmax} и q_{nmin} - формируются субъективно разработчиком исходя из условия задач. Действительно, как выбирать, как назначать какие-либо весовые коэффициенты, пока не определился с тем, что необходимо. Исходя из этого посыла предполагаем, что задать подобные значения q_{nmax} и q_{nmin} возможно всегда. А теперь очень важный момент, следующий из представленной выше интерпретации. Задав q_{nmax} и q_{nmin} , мы тем самым задаем идеальный вектор, причем не только его величину, но и расположение в пространстве. Напомним, что, следуя интерпретации, расположение вектора в

пространстве задается назначением весовых коэффициентов, те. В данном случае, задав q_{nmax} и q_{nmin}, мы тем самым задаем и весовые коэффициенты – относительную важность параметров. Т.е. при таком подходе к решению задачи снимаются все неопределенности – выбирается вариант, наиболее подходящий сформулированным требованиям.

Замечание:

В случае если у варианта значение какого-либо параметра превосходит требуемое, то при выборе учитывается требуемое качество, а не значение какого-либо параметра варианта (так как в подобном завышении локального качества нет необходимости). Усеет же завышенного значения параметра приведет к изменению расположения идеального вектора в пространстве, т.е. к изменению «весов», задаваемых по умолчанию.

Для решения задачи используется аддитивная свертка:

$$\overline{Yn} = \sum_{n=1}^{N} \overline{q}_{nm}$$

а нормированные значения параметров вариантов определяются соответствующим образом:

$\overline{q}_{nm} = q_{nm} / q_{nmax}$, если лучшее значение качества определяется, как max (например, производительность)

$\overline{q}_{nm} = q_{nmin} / q_{nm}$, если лучшее значение качества определяется, как min (например, стоимость)

Достоинства:

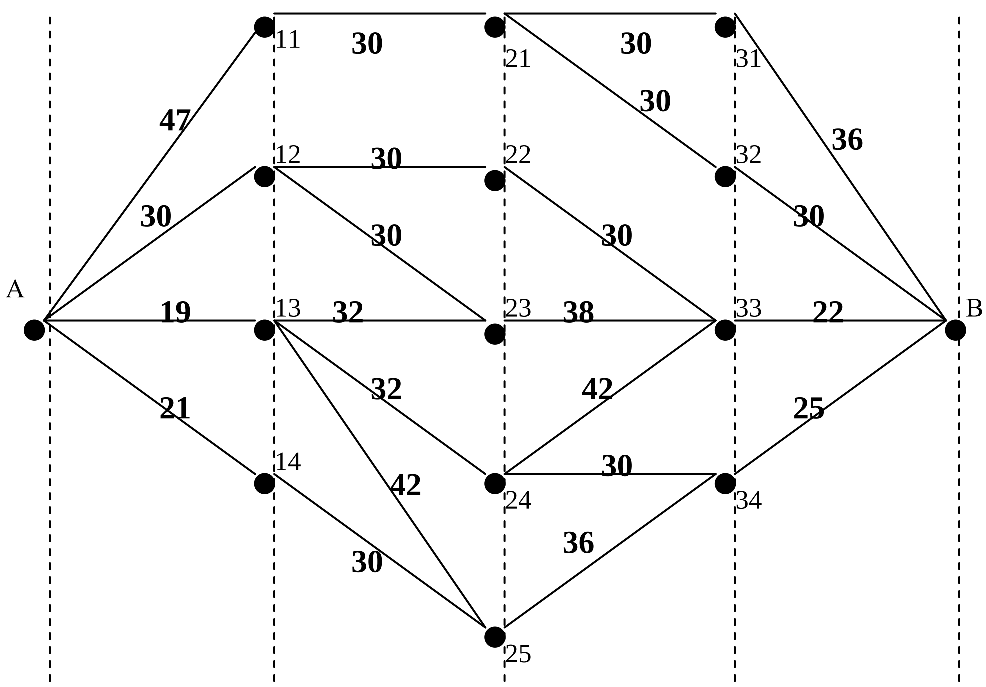
- Однозначно интерпретируем идеальный вектор качества, задающий то совокупное качество, которое следует обеспечить проектируемой системой (а не неким гипотетически идеальным набором значений параметров, достижение которых может быть недостижимым в совокупности).
- На идеальный вектор (величину и расположения) никак не сказывается выбор сравниваемых между собою вариантов.
- Не требуется как-либо дополнительно учитывать относительную важность параметров, так как она подсознательно задается разработчиком при формировании требуемого качества системы при проектировании.

Другими словами, вся эвристика здесь сводится к определению того, что требуется, а без этого (без формулирования цели операции) невозможно говорить об использовании методов оптимизации.

Ограничения на использование векторной интерпретации оптимизационных задач. Динамическое программирование.

Методы векторной оптимизации имеют массу практических приложений в нашей области знаний, однако существуют и ограничения на их использование. Они могут использоваться в том случае, когда векторы качества вариантов решений совпадают для всех этапов решения. Рассмотрим одно из исключений – задачу выбора оптимального маршрута, решаемую методом динамического программирования.

Задачей при разработке метода динамического программирования являлось уменьшения вычислительной сложности решения, за счет отказа от полного перебора вариантов.



Рассмотрим пример постановки оптимизационной задачи: пусть требуется проанализировать сеть связи, где сообщения должно попасть из А в В. Возможные маршруты и расстояния (в качестве критериев могут использоваться и любые иные параметры, по которым необходимо произвести выбор, например, скорость доставки сообщения, вероятность, цена доставки и т.п.) указаны на рисунке. Требуется найти наикротчайший путь (маршрут).

Принятие очередного решения или выбор маршрута равносилен выбору направления выхода из того узла коммутации (УК), куда попало сообщение.

Идея метода динамического программирования.

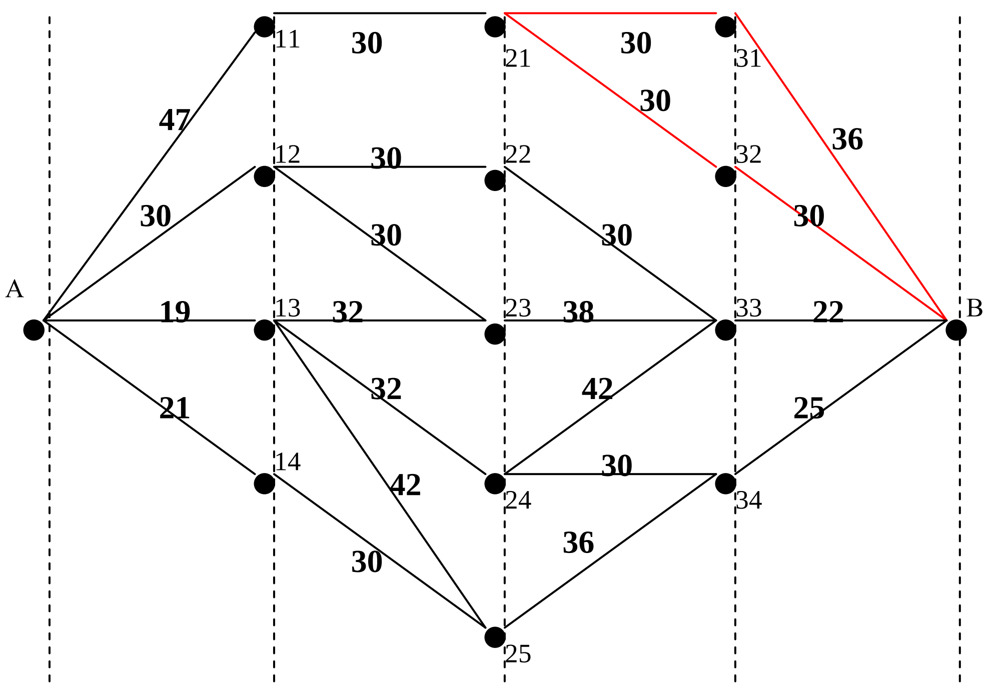
Возможный подход к решению задачи состоит в полном переборе возможных решений. Т.е. могут быть рассчитаны длины всех путей и выбран минимальный из них. Метод же динамического программирования предполагает построение оптимального решения шаг за шагом на каждом этапе оптимизируя только один шаг. Здесь лежит общая идея данного класса методов, состоящая в том, что всегда алгоритмически проще многократно решать простую задачу, чем однократно – сложную. Проблема задач динамического программирования состоит в том, что оптимизация маршрутов вне зависимости от выбора оптимальных решений на последующих или предыдущих шагах, невозможна, т.к. выигрыш, полученный на отдельном шаге может не дать совокупного выигрыша с учетом всех выбранных шагов маршрута, т.е. оптимальный маршрут может быть не найден. Следовательно, планируя многошаговую операцию, на отдельном шаге в качестве оптимального надо выбирать решение, с учетом всех его последствий (как на предшествующих, так и на последующих шагах).

Метод динамического программирования состоит в следующем: планируя некий m-й шаг, делают различные предположения о том, чем закончился предыдущий (m-1), и для каждого из этих предположений находятся так называемые условно-оптимальные решения на m-ом шаге (условно-оптимальное, т.к. данное оптимальное решение находится в предположении или при условии, что m-1 шаг закончился так-то). Аналогичным образом могут быть последовательно выбраны оптимальные шаги на всех шагах.

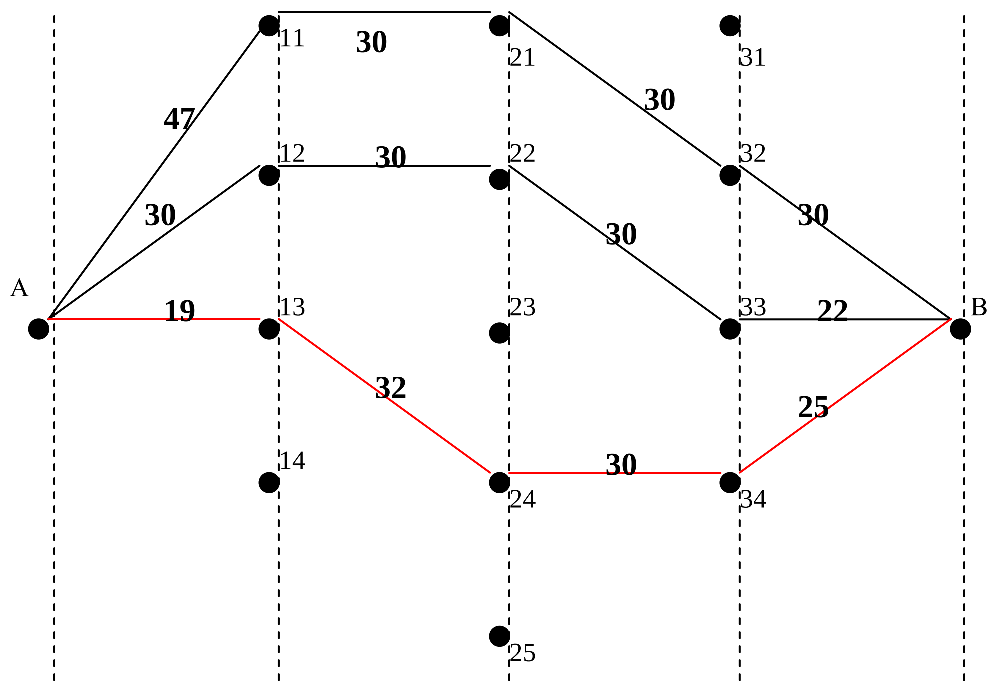
В процессе оптимизации методом динамического программирования многошаговый процесс осуществляется дважды:

1. От конца к началу, с целью выявления условно-оптимальных решений.
2. От начала к концу, с целью определения оптимальных маршрутов.

А собственно интерпретация (или идея метода решения) весьма проста. Если на промежуточных шагах существует несколько вариантов прохождения одного и того же пути (между одними и теми же УК) несколькими маршрутами, можно сравнить эти маршруты и оставить для последующего анализа оптимальный.
Вернемся к примеру (см. рисунок) и проиллюстрируем идею.



Следуя из точки В в обратном направлении, в предположении, что мы оказались в узле с номером 21, видим, что из УК21 существует два маршрута в точку В (выделены красным цветом), т.е. мы можем сравнить эти два варианта (выделены красным) между собой, выбрав из них условно-оптимальное решение. Критерий одного варианта определяется, как 30+36 = 66, другого 30+30 = 60 (это условно оптимальное решение), другое можно вычеркнуть. Другими словами, этим мы определяем, что при условии попадания сообщения в УК21, далее оно должно следовать по выбранному нами оптимальному маршруту.
Пройдем аналогично из точки В по всем направлением, получаем преобразованный вариант оптимизационной задачи, состоящий из условно-оптимальных решений, см. рисунок.



Теперь остается пройти из точки А в точку В по условно-оптимальным маршрутам, рассчитать для них значения критерия оптимальности - выбрать оптимальный.

Оптимальным будет маршрут, выделенный красным цветом на рисунке. Для его определения не потребовалось осуществлять полный перебор возможных маршрутов с расчетом для них критерия оптимальности с последующим сравнением всех возможных вариантов.