ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»
Кафедра «Техносферная и информационная безопасность»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Б1.В.ДВ.07.01 Теория принятия решений

Направление подготовки: 20.03.01 «Техносферная безопасность»

Профиль подготовки: «Безопасность жизнедеятельности в техносфере»

Квалификация (степень) выпускника: бакалавр

Нормативный срок обучения: 4 года

Форма обучения: очная

СОДЕРЖАНИЕ

1. КОНСПЕКТ ЛЕКЦИЙ

1.1 Лекция № 1
Тема: Основные понятия и принципы исследования операций

1.1.1 Вопросы лекции:

1. История возникновения понятий «системный анализ». «Исследование операций», состав дисциплин входящих в «Исследование операций» и «системный анализ».

2. Основные принципы системного анализа и теории принятия решений.

1.1.2 Краткое содержание вопросов:

1. История развития системного анализа

Составляющим понятий «системный анализ», «системная проблема», «системное исследование» является слово «система», которое появилось в Древней Элладе 2000—2500 лет назад и первоначально означало: сочетание, организм, устройство, организация, строй, союз. Оно также выражало определенные акты деятельности и их результаты (нечто, поставленное вместе; нечто, приведенное в порядок).

Первоначально слово «система» было связано с формами социально-исторического бытия. Лишь позднее принцип порядка, идея упорядочения переносятся на Вселенную.

Перенос значения слова с одного объекта на другой и вместе с тем превращение слова в обобщенное понятие совершаются поэтапно. Метафоризация слова «система» была начата Демокритом (460—360 до н. э.), древнегреческим философом, одним из основоположников материалистического атомизма. Образование сложных тел из атомов он уподобляет образованию слов из слогов и слогов из букв. Сравнение неделимых форм (элементов с буквами) — один из первых этапов формирования научно-философского понятия, обладающего обобщенным универсальным значением.

На следующем этапе происходят дальнейшая универсализация значения слова, наделение его высшим обобщенным смыслом, что позволяет применять его и к физическим, и к искусственным объектам. Универсализация может осуществляться двояко — или в процессе мифотворчества, т. е. построения мифа на основе метафоры характерно для одного из основателей объективного идеализма Платона (427—347 до н. э.), или же путем воссоздания философско-рациональной картины мироздания и человеческой культуры, т. е. трансформирования и развертывания метафоры в философской системе характерно для Аристотеля (384—322 до н. э.), колеблющегося между материализмом и идеализмом.

Итак, в античной (древней) философии термин «система» характеризовал упорядоченность и целостность естественных объектов, а термин «синтагма» — упорядоченность и целостность искусственных объектов, прежде всего продуктов познавательной деятельности. Именно в этот период был сформулирован тезис о том, что целое больше суммы его частей (Философский словарь. М.: Политиздат, 1980).

Не касаясь вопроса о трактовке системности знания в средневековой философии, отметим лишь, что для выражения интегративности познавательных образований здесь стали использоваться новые термины: сумма, дисциплина, доктрина...

С возникновением науки и философии Возрождения (XV в.) связано радикальное преобразование в истолковании бытия. Трактовка бытия как космоса сменяется рассмотрением его как системы мира. При этом система мира понимается как независимое от человека, обладающее своим типом организации, иерархией, имманентными (свойственными, внутренне присущими какому-либо предмету, явлению, проистекающими из их природы) законами и суверенной структурой. Кроме того, бытие становится не только предметом философского размышления, стремящегося постичь его целостность, но и предметом социально-научного анализа. Возникает ряд научных дисциплин, каждая из которых вычленяет в природном мире определенную область и анализирует ее свойственными этим дисциплинам методами.

Астрономия была одной из первых наук, которая перешла к онтолого-натуралистической интерпретации системности мироздания. Большую роль в становлении новой трактовки системности бытия сыграло открытие Н. Коперника (1473—1543). Он создал Гелиоцентрическую систему мира, объяснив, что Земля, как и другие планеты, обращается вокруг Солнца и, кроме того, вращается вокруг своей оси. Телеологизм (телеология — учение о совершенстве, учение о конечных причинах — воззрение, объясняющее закономерную связь явлений природы не объективными причинами, а целями, установленными Божьей волей), отягощавший представления Коперника, был преодолен позднее Г. Галилеем (1564—1642) и И. Ньютоном (1642—1727).

Наука эпохи Возрождения выработала определенную концептуальную систему. Ее важнейшие категории — вещь и свойства, целое и часть, субстанция и атрибуты. Вещь трактовалась как сумма отдельных свойств.

Основная познавательная процедура сводится к поиску сходства и различия в предметах. В связи с этим весьма специфично трактуется категория «отношение», которая выражает прежде всего субординацию главных и второстепенных свойств, динамическое воздействие некоего предмета на другой, первый из которых является причиной, а второй — следствием.

Важнейшая особенность представлений о системности предмета познания, характерная для науки эпохи Возрождения, состоит в выдвижении на первый план каузального, а не телеологического способа объяснения...

Глубокую и основательную разработку идея системной организации научного знания получила в немецкой классической философии. Структура научного знания, принципы и основания построения теоретических систем стали в ней предметом специального философского, логико-методологического анализа.

Немецкий математик и философ И.Г. Ламберт (1728—1777) подчеркивал, что «всякая наука, как и ее часть, предстает как система, поскольку система есть совокупность идей и принципов, которая может трактоваться как целое. В системе должны быть субординация и координация». Следует отметить, что он анализировал системность науки на основе обобщенного рассмотрения систем вообще, построения общей системологии.

Новый этап в интерпретации системности научного знания связан с именем И. Канта (1724—1804). Его заслуга состоит не только в четко осознанном системном характере научно-теоретического знания, но и в превращении этой проблемы в методологическую, в выявлении определенных процедур и средств системного конструирования знания.

Ограниченность кантовского понимания системности знания состоит в том, что конструктивно-методологические принципы образования научных систем являются у него характеристиками лишь формы, а не содержания знания.

Эту линию в еще большей мере проводит И.Г. Фихте (1762—1814), который считает, что принципы полагания формы знания являются одновременно принципами полагания и его содержания. Исходный тезис Фихте — научное знание есть системное целое. Фихте является родоначальником того направления в классической немецкой философии, которое останавливается на вычленении формально-логических принципов систематизации сложившегося знания, ограничивая тем самым системность знания систематичностью его формы. Это привело к отождествлению системности научного знания и его систематического изложения. Это направление сосредоточивает свое внимание не на научном исследовании, а на изложении результатов знания, систематического представления теоретического знания. Такой подход особенно проявился у последователей Канта и Фихте — К. Шмида, Я. Фриза и др.

Г. Гегель (1770—1831), объективный идеалист, исходит из единства содержания и формы знания, из тождества мысли и действительности и предлагает историческую трактовку становления системы в соответствии с принципом восхождения от абстрактного к конкретному. Однако в силу отождествления метода и системы, в силу телеологического истолкования истории знания, он не смог предложить методолого-конструктивных средств для формирования системных научных образований и фактически лишил все предшествующие ему теоретические и философские построения статуса системы. По сути дела, они оказались в его интерпретации лишь абстрактным выражением, превращенной формой его системы, претендовавшей на единственно возможную и абсолютно значимую.

Теоретическое естествознание XIX—XX вв. исходит из различения предмета и объекта знания. Подчеркивая активный характер человеческого познания, новый способ мысли трактует предмет исследований как нечто созданное и создаваемое человеком в ходе освоения природы. Поднимается роль моделей в познании.

Целое понимается уже не как простая сумма, а как функциональная совокупность, которая формируется некоторым заранее задаваемым отношением между элементами. При этом фиксируется наличие особых интегративных характеристик данной совокупности — целостность, несводимость к составляющим элементам. Сама эта совокупность, отношение между элементами (их координация, субординация и т.д.) определяются некоторым правилом или системообразующим принципом. Этот принцип относится как к порождению свойств целого из элементов, так и к порождению свойств элементов из целого. Системообразующий принцип позволяет не только постулировать те или иные свойства элементов и системы, но и предсказывать возможные элементы и свойства системной совокупности.

Марксистская интерпретация системности научного знания противостоит как наивному антологизму, так и волюнтаристскому конструктивизму. В противовес созерцательному материализму марксизм подчеркивает активный характер человеческого познания, связывает системность научного знания с формами познавательной деятельности человека. Вместе с тем марксистское понимание познания как деятельности не имеет ничего общего с волюнтаристской ее трактовкой, лишающей мышление содержательных характеристик. Марксизм подчеркивает единство природы и деятельности человека, проводит мысль о том, что «человек в процессе производства может действовать лишь так, как действует сама природа, т.е. может изменять лишь формы веществ».

Марксистская гносеология выдвинула определенные принципы анализа системности научного знания. К ним относятся историзм, единство содержательной и формальной сторон научного знания, трактовка системности не как замкнутой системы, а как развивающейся последовательности понятий и теорий. При таком подходе системность знаний предполагает дальнейшее совершенствование системы понятий...

Попытки разработать общие принципы системного подхода были предприняты врачом, философом и экономистом А.А. Богдановым (1873— 1928) в работе «Всеобщая организационная наука (тектология)» (3-е изд. М.; Л., 1925—1929. Ч. 1—3). Исследования, проведенные уже в наши дни, показали, что важные идеи и принципы кибернетики, сформулированные Н. Винером и особенно У. Росс Эшби, значительно раньше, хотя и в несколько иной форме, были выражены Богдановым. В еще большей мере это относится к общей теории систем (ОТС) Л. фон Берталанфи, идейная часть которой во многом предвосхищена автором тектологии.

Тектология (греч. — «строитель») — весьма оригинальная общенаучная концепция, исторически первый развернутый вариант ОТС. Ее созданием автор хотел бросить вызов марксизму, выдвинув в противовес ему концепцию, которая претендует на универсальность. Для построения тектологии используется материал самых различных наук, в первую очередь естественных. Анализ этого материала приводит к выводу о существовании единых структурных связей и закономерностей, общих для самых разнородных явлений.

Основная идея тектологии — признание необходимости подхода к любому явлению со стороны его организованности (у других авторов — системности). Под организованностью понимается свойство целого быть больше суммы своих частей. Чем больше целое разнится от суммы своих частей, тем больше оно организовано. Тектология рассматривает все явления как непрерывные процессы организации и дезорганизации. Принципы организованности и динамичности тесно связаны с принципом целостного рассмотрения отдельных явлений и всего мира вообще.

ОТС и тектология — это две науки об организованности, системности явлений, кибернетика же — наука об управлении этими объектами. Таким образом, предмет кибернетики уже, что обусловлено большей широтой понятия «организация системы», чем понятия «управление». Тектология как общая теория включает в сферу своего внимания не только кибернетические принципы, т. е. принципы управления систем, но и вопросы их субординации (иерархических порядков), их распада и возникновения, обмена со средой и веществом и т.д.

Австрийский биолог и философ Л. Фон Берталанфи (1901—1972) первым из западных ученых разработал концепцию организма как открытой системы и сформулировал программу построения ОТС. В своей теории он обобщил принципы целостности, организации, эквифинальности (достижения системой одного и того же конечного состояния при различных начальных условиях) и изоморфизма.

Начиная со своих первых работ, Л. Берталанфи проводит мысль о неразрывности естественно-научного (биологического) и философского (методологического) исследований... Сначала была создана теория открытых систем, граничащая с современной физикой, химией и биологией. Классическая термодинамика исследовала лишь закрытые системы, т. е. не обменивающиеся веществом с внешней средой и имеющие обратимый характер. Попытка применения классической термодинамики к живым организмам (начало XX в.) показала, что, хотя при рассмотрении органических явлений использование физико-химических принципов имеет большое знание, так как в организме имеются системы, находящиеся в равновесии (характеризующимся минимумом свободной энергии и максимумом энтропии), однако сам организм не может рассматриваться как закрытая система в состоянии равновесия, ибо он не является таковым. Организм представляет собой открытую систему, остающуюся постоянной при непрерывном изменении входящих в нее веществ и энергии (так называемое состояние подвижного равновесия).

В 1940—50 гг. Л. Берталанфи обобщил идеи, содержащиеся в теории открытых систем, и выдвинул программу построения ОТС, являющейся всеобщей теорией организации. Проблемы организации, целостности, направленности, телеологии, саморегуляции, динамического взаимодействия весьма актуальны и для современной физики, химии, физической химии и технологии, а не только для биологии, где подобные проблемы встречаются повсюду. Пока что такие понятия были чужды классической физике. Если до сих пор унификацию наук видели обычно в сведении всех наук к физике, то, с точки зрения Л, Берталанфи, единая концепция мира может быть, скорее, основана на изоморфизме законов в различных областях. В результате он приходит к концепции синтеза наук, которую и противоположность редукционизму (т. е. сведению всех наук к физике) называет перспективизмом.

Построенная теория организации является специальной научной дисциплиной. Вместе с тем она выполняет определенную методологическую функцию. В силу общего характера исследуемого предмета (системы) ОТС дает возможность охватить одним формальным аппаратом обширный круг специальных систем. Благодаря этому она может освободить ученых от массового дублирования работ, экономя астрономические суммы денег и времени.

К числу недостатков ОТС Л. Берталанфи относятся неполное определение понятия «система», отсутствие особенностей саморазвивающихся систем и теоретического исследования связи, а также условий, при которых система модифицирует свои формы. Но основной методологический недостаток его теории заключается в утверждении автора о том, что она выполняет роль философии современной науки, формируя философски обобщенные принципы и методы научного исследования. В действительности это не так. Ибо для философского учения о методах исследования необходимы совершенно иные (новые) исходные понятия и иная направленность анализа: абстрактное и конкретное специфически мысленное знание, связь знаний, аксиоматическое построение знаний и др., что отсутствует в ОТС.

Однако, учитывая большое методологическое значение работы Л. Берталанфи (Общая теория систем — обзор проблем и результатов. Системные исследования // Ежегодник. М.: Наука, 1969), рассмотрим различные направления в разработке теории систем. В соответствии с его взглядами, системная проблематика сводится к ограничению применения традиционных аналитических процедур в науке. Обычно системные проблемы выражаются в полу-метафизических понятиях и высказываниях, подобных, например, понятию «эмерджентная эволюция» или утверждению «целое больше суммы его частей», однако они имеют вполне определенное операционное значения. При применении «аналитической процедуры» некоторая исследуемая сущность разлагается на части, и, следовательно, затем она может быть оставлена или воссоздана из собранных вместе частей, причем эти процессы возможны как мысленно, так и материально. Это основной принцип «классической» науки, который может осуществляться различными путями: разложением исследуемого явления на отдельные причинные цепи, поисками «атомарных» единиц в различных областях науки и т. д. Научный прогресс показывает, что этот принцип классической науки, впервые сформулированный Галилеем и Декартом, приводит к большим успехам при изучении широкой сферы явлений.

Применение аналитических процедур требует выполнения двух условий. Во-первых, необходимо, чтобы взаимодействие между частями данного явления отсутствовало или было бы пренебрежимо мало для некоторой исследовательской цели. Только при этом условии части можно реально, логически или математически «извлекать» из целого, а затем «собирать». Во-вторых, отношения, описывающие поведение частей, должны быть линейными. Только в этом случае имеет место отношение суммативности, т. е. форма уравнения, описывающего поведение целого, такова же, как и форма уравнений, описывающих поведение частей; наложение друг на друга частных процессов позволяет получить процесс в целом и т.д.

Для образований, называемых системами, т.е. состоящих из взаимодействующих частей, эти условия не выполняются. Прототипом описания систем являются системы дифференциальных уравнений, в общем случае нелинейных. Систему, или «организованную сложность», можно описать через «сильные взаимодействия» или взаимодействия, которые «нетривиальны», т.е. нелинейны. Методологическая задача теории систем, таким образом, состоит в решении проблем, которые носят более общий характер, чем аналитически-суммативные проблемы классической науки.

Существуют различные подходы к таким проблемам. Автор намеренно использует довольно расплывчатое выражение — «подходы», поскольку они логически неоднородны, характеризуются различными концептуальными моделями, математическими средствами, исходными позициями и т.д. Однако все они являются теориями систем. Если оставить в стороне подходы в прикладных системных наследованиях, таких как системотехника, исследование операций, линейное и нелинейное программирование и т.д., то наиболее важными являются следующие подходы:

- «Классическая» теория систем. Эта теория использует классическую математику и имеет цели: установить принципы, применимые к системам вообще или к их определенным подклассам (например, к закрытым и открытым системам); разработать средства для их исследования и описания и применить эти средства к конкретным случаям. Учитывая достаточную общность получаемых результатов, можно утверждать, что некоторые формальные системные свойства относятся к любой сущности, которая является системой (к открытым системам, иерархическим системам и т.д.), даже если ее особая природа, части, отношения и т.д., не известны или не исследованы. Примерами могут служить: обобщенные принципы кинетики, применимые, в частности, к популяциям молекул или биологических существ, т.е. к химическим и биологическим системам; уравнения диффузии, используемые в физической химии и для анализа распространения слухов; понятия устойчивого равновесия и модели статистической механики, применимые к транспортным потокам; аллометрический анализ биологических и социальных систем.
- Использование вычислительных машин и моделирование. Системы дифференциальных уравнений, применяемые для «моделирования» или спецификации систем, обычно требуют много времени для решения, даже если они линейны и содержат немного переменных; нелинейные системы уравнений разрешимы только в некоторых частных случаях. По этой причине с использованием вычислительных машин открылся новый подход к системным исследованиям. Дело не только в значительном облегчении необходимых вычислений, которые иначе потребовали бы недопустимых затрат времени и энергии, и замене математической изобретательности заранее установленными последовательностями операций. Важно еще и то, что при этом открывается доступ в такие области, где в настоящее время отсутствует соответствующая математическая теория и нет удовлетворительных способов решения. Так, с помощью вычислительных машин могут анализировать системы, по своей сложности далеко превосходящие возможности традиционной математики; с другой стороны, вместо лабораторного эксперимента можно воспользоваться моделированием на вычислительной машине и построенная таким образом модель затем может быть проверена в реальном эксперименте. Таким способом Б. Гесс, например, рассчитал 14-звенную цепь реакций гликолиза в клетке на модели, содержащей более 100 нелинейных дифференциальных уравнений. Подобный анализ стал обычным делом в экономических разработках, при исследовании рынка и т. д.
- Теория ячеек. Одним из аспектов системных исследований, который следует выделить, поскольку эта область разработана чрезвычайно подробно, является теория ячеек, изучающая системы, составленные из подъедениц с определенными граничными условиями, причем между этими подъединицами имеют место

процессы переноса. Такие ячеечные системы могут иметь, например, «цепную» или «сосковую» структуру (цепь ячеек или центральную ячейку, сообщающуюся с рядом периферийных ячеек). Вполне понятно, что при наличии в системе трех и более ячеек математические трудности становятся чрезвычайно большими. В этом случае анализ возможен лишь благодаря использованию преобразований Лапласа и аппарата теорий сетей и графов.

- Теория множеств. Общие формальные свойства систем и формальные свойства закрытых и открытых систем могут быть аксиоматизированы в языке теории множеств. По математическому изяществу этот подход выгодно отличается от более грубых и специализированных формулировок «классической» теории систем. Связи аксиоматизированной теории систем с реальной проблематикой системных исследований пока выявлены весьма слабо.
- Теория графов. Многие системные проблемы относятся к структурным и топологическим свойствам систем, а не к их количественным отношениям. В этом случае используется несколько различных подходов. В теории графов, особенно в теории ориентированных графов (диграфов), изучаются реляционные структуры, представляемые в топологическом пространстве. Эта теория применяется для исследования реляционных аспектов биологии. В магматическом смысле она связана с матричной алгеброй, но своими моделями — с тем разделом теории ячеек, в котором рассматриваются системы, содержащие частично «проницаемые» подсистемы, а вследствие этого — с теорией открытых систем.
- Теория сетей. Эта теория, в свою очередь, связана с теориями множеств, графов, ячеек и т. д. Она применяется к анализу таких систем, как нервные сети.
- Кибернетика. В основе кибернетики, т.е. теории систем управления, лежит связь (передача информации) между системой и средой и внутри системы, а также управление (обратная связь) функциями системы относительно среды. Кибернетические модели допускают широкое применение, но их нельзя отождествлять с теорией систем вообще. В биологии и других фундаментальных науках кибернетические модели позволяют описывать формальную структуру механизмов регуляции, например, при помощи блок-схем и графов потоков. Использование кибернетических моделей позволяет установить структуру регуляции системы даже в том случае, когда реальные механизмы остаются неизвестными и система представляет собой «черный ящик», определяемый только его входом и выходом. Таким образом, одна и та же кибернетическая схема может применяться к гидравлическим, электрическим, физиологическим и другим системам. Тщательно разработанная техническая теория сервомеханизмов применяется естественным системам в ограниченном объеме.
- Теория информации. По К. Шеннону, математическое выражение для понятия информации изоморфно выражению для негэнтропии в термодинамике. Считается, что понятие информации можно использовать в качестве меры организации. Хотя теория информации имеет большое значение для техники связи, ее применение в науке весьма незначительно. Главной проблемой остается выяснение отношения между информацией и организацией, между теорией информации и термодинамикой.
- Теория автоматов. Это так называемая теория абстрактных автоматов, имеющих вход, выход, иногда способных действовать методом проб и ошибок и обучаться. Общей моделью теории автоматов является машина Тьюринга, которая представляет собой абстрактную машину, способную печатать (или стирать) на ленте конечной длины цифры 1 и 0. Можно показать, что любой сколь угодно сложный процесс может моделироваться машиной Тьюринга, если этот процесс можно выразить конечным числом операций. В свою очередь, то, что возможно логически (т.е. в алгоритмическом символизме), может также быть

сконструировано — в принципе, но не всегда практически — автоматом (т, е. алгоритмической машиной).

- Теория игр. Несмотря на то, что теория игр несколько отличается от других рассмотренных системных подходов, все же ее можно поставить в ряд наук о системах. Н ней рассматривается поведение «рациональных» игроков, пытающихся достичь максимальных выигрышей и минимальных потерь за счет применения соответствующих стратегий в игре с соперником (или природой). Следовательно, теория игр рассматривает системы, включающие антагонистические силы.
- Теория решений. Эта математическая теория изучает условия выбора между альтернативными возможностями.
- Теория очередей. Рассматривает оптимизацию обслуживания при массовых запросах.

Несмотря на неоднородность и явную неполноту проведенного рассмотрения, отсутствие достаточной четкости в различении моделей (например, моделей открытой системы, цепи обратной связи) и математических формализмов (например, формализмов теорий множеств, графов, игр), такое перечисление позволяет показать, что существует целый ряд подходов к исследованию систем, а некоторые из них обладают мощными математическими методами. Проведение системных исследований означает прогресс в анализе проблем, которые ранее не изучались, считались выходящими за пределы науки или чисто философскими.

Хорошо известно, что проблема соответствия между моделью и реальностью чрезвычайно сложна. Нередко мы располагаем тщательно разработанными математическими моделями, но остается неясным, как можно применять их в конкретном случае. Для многих фундаментальных проблем вообще отсутствуют подходящие математические средства. Чрезмерные ожидания привели в последнее время к разочарованию. Так, кибернетика продемонстрировала свое влияние не только в технике, но и в фундаментальных науках; построила модели ряда конкретных явлений, показала научную правомерность телеологического объяснения и т.д. Тем не менее кибернетика не создала нового широкого «мировоззрения», оставаясь скорее расширением, чем заменой механистической концепции. Теория информации, математические основы которой детально разработаны, не смогла построить интересных приложений в психологии и социологии. Большие надежды возлагались на применение теории игр к вопросам войны и политики, но едва ли можно считать, что она улучшила политические решения и положение дел в мире. Эту неудачу можно было ожидать, учитывая, как мало существующие державы походят на «рациональных» игроков теории игр. Понятия и модели равновесия, гомеостазиса, регулирования приложимы для описания процессов функционирования систем, но они неадекватны для анализа явлений измерения, дифференциации, эволюции, уменьшения энтропии, творчества и т.д. Это осознавал Кэннон, когда допускал кроме гомеостазиса еще и гетеростазис, характеризующий такие явления. Теория открытых систем широко применяется для описания явлений биологии (и техники), но необходимо предостеречь против неосмотрительного распространения ее на те области, для которых она не предназначена. Вполне очевидно, что отмеченные ограниченности системных научных подходов, существующих едва ли больше двадцати-тридцати лет, совершенно естественны. В конечном счете разочарование, о котором мы только что говорили, объясняется применением моделей, полезных в определенных аспектах, к проблемам метафизического и философского порядка.

Несмотря на то что математические модели обладают важными достоинствами — четкостью, возможностью строгой дедукции, проверяемостью и т.д., — не следует отказываться от использования моделей, сформулированных в обычном языке.

Вербальная модель лучше, чем отсутствие модели вообще или математическая модель, которая при насильственном насаждении фальсифицирует реальность. Многие теории, получившие огромное влияние в науке, являются не математическими по своему характеру (например, психоаналитическая теория), а в других случаях лежащие и их основе математические конструкции осознаются позднее и охватывают лишь отдельные аспекты соответствующих эмпирических данных (как в теории отбора).

Математика, по сути дела, сводится к установлению (алгоритмов, которые более точны, чем алгоритмы обычного языка. История науки свидетельствует о том, что описание проблем на обычном языке часто предшествует их математической формулировке, т.е. отысканию алгоритма. Приведем несколько хорошо известных примеров: знаки, используемые для обозначения чисел и счета, эволюционировали от слов естественного языка к римским цифрам (полувербальным, несовершенным, полуалгебраическим) и далее — к арабской численной символике, в которой важное значение имеет положение знака; уравнения первоначально формулировались в словесной форме, затем — с использованием примитивного символизма, который мастерски применял Диофант и другие основатели алгебры, и, наконец, в современном символизме; для многих теорий, например для теории Дарвина, математические основы определяются значительно позднее, чем создаются. Вероятно, лучше иметь сначала какую-то нематематическую модель со всеми ее недостатками, но охватывающую некоторый не замеченный ранее аспект исследуемой реальности и позволяющую надеяться на последующую разработку соответствующего алгоритма, чем начинать со скороспелых математических моделей.

Таким образом, модели, выраженные в обычном языке, оставляют себе место в теории систем. Идея системы сохраняет значение даже там, где ее нельзя сформулировать математически или где она остается скорее направляющей идеей, чем математической конструкцией. Например, у нас может не быть удовлетворительных системных понятий для социологии; однако само понимание того, что социальные сущности являются системами, а не суммами социальных атомов, или того, что история имеет дело с системами {хотя бы и плохо определенными), называемыми цивилизациями, которые подчиняются общим для систем принципам, подразумевает важную переориентацию в рассматриваемых научных областях.

Как мы видели ранее, в рамках системного подхода существуют и механистические, и организмические тенденции и модели, пытающиеся познать системы либо с помощью таких понятий, как «анализ», «линейная (включая круговую) причинность», «автомат» и т.д., либо при помощи понятий «целостность», «взаимодействие», «динамика» и им подобных. Эти два типа моделей не исключают друг друга и даже могут использоваться для описания одних и тех же явлений.

Итак, подводя итоги, ОТС у Л. Берталанфи выступает в двух смыслах. В широком — как основополагающая, фундаментальная наука, охватывающая всю совокупность проблем, связанных с исследованием и конструированием систем. В теоретическую часть включаются 12 направлений, приведенных выше. В узком смысле — ОТС, стремящаяся вывести из общего определения системы как комплекса взаимодействующих элементов понятия, относящиеся к организованным целым (взаимодействие, сумма, централизация, фатальность и т.д.), и применяющая их к анализу конкретных явлений. Прикладная область общей теории систем включает, согласно Берталанфи: 1) системотехнику; 2) исследование операций; 3) инженерную психологию (схема 1.1).

Системные исследования — вся совокупность научных и технических проблем, которые при всей их специфике и разнообразии сходны в понимании и рассмотрении исследуемых ими объектов как систем, т. е. множества взаимосвязанных элементов, выступающих в виде единого целого.

Соответственно этому системный подход — эксплицитное (разъяснительное) выражение процедур представления объектов как систем и способов их описания, объяснения, предвидения, конструирования и т. д.

Общая теория систем, таким образом, выступает в этом случае как обширный комплекс научных дисциплин. Следует, однако, отметить, что при таком истолковании в известной мере теряется определенность задач теории систем и ее содержания. Строго научной концепцией (с соответствующим аппаратом, средствами и т.д.) можно считать лишь общую теорию систем в узком смысле. Что же касается общей теории систем в широком смысле, то она или совпадает с общей теорией систем в узком смысле (один аппарат, одни исследовательские средства и т.д.), или представляет собой действительное расширение и обобщение общей теории систем в узком смысле и аналогичных дисциплин, однако тогда встает вопрос о развернутом представлении ее средств, методов, аппарата и т.д. Без ответа на этот вопрос общая теория систем в широком смысле фактически остается лишь некоторым проектом (пусть даже очень заманчивым) и вряд ли может быть развита в строгую научную теорию.

Общая теория систем в широком смысле (по Берталанфи) — фундаментальная наука, охватывающая всю совокупность проблем, связанных с исследованием и конструированием систем

Теоретическая часть **Прикладная область**

1. Кибернетика — базируется на принципе обратной связи и круговых причинных целях и исследует механизмы целенаправленного и самокотролируемого поведения; теория систем управления

2. Теория информации, вводящая понятие количества информации и развивающая принципы передачи информации

3. Теория игр — рассматривает поведение игроков, пытающихся достичь максимального выигрыша и минимальных потерь за счет применения соответствующих стратегий в игре с соперником

4. Теория решений — математический теория, изучающая условия выбора между альтернативными возможностями

5. Топология, включающая теорию сетей и теорию графов

6. Факториальный анализ

7. ОТС в узком смысле, которая стремится вывести из общего

1. Системотехника — направление в кибернетике, изучающее вопросы планирования, проектирования и поведения сложных систем различного назначения (АСУ, человеко-машинные комплексы и др.), при котором составляющие системы рассматриваются во взаимодействии, несмотря на их разнородность. Основным методом системотехники является системный анализ. Центральное техническое звено комплекса — ЭВМ, человеческое звено — оператор. Системотехника играет важную роль в развитии инженерной психологии, так как для проектирования комплексов необходимо учитывать характеристики человека

2. Исследование операций — изучает прикладное направление кибернетики, использующее математические методы для обоснования решения во всех областях человеческой деятельности

3. Инженерная психология — отрасль психологии, исследующая процессы и средства информационного взаимодействия между человеком и машиной. Инженерная психология возникла в условиях научно-технической революции, преобразовавшей психологическую структуру

определения системы как комплекса взаимодействующих элементов, понятий, относящихся к организованным целым (взаимодействие, сумма, фатальность, централизация и т.д.) и применение их к анализу конкретных явлений	производственного труда, важнейшими составляющими которого стали восприятие и переработка оперативной информации, принятие решений в условиях ограниченного времени

Схема 1.1 — Состав общей теории систем

Системное движение по своим задачам действительно призвано выработать новое — в противовес механистическому — видение мира, разработать принципы нового направления научных и технических исследований. И как таковое оно, несомненно, должно включать в себя совокупность принципиально различных по своему типу разработок — философских, логико-методологических, математических, модельных, эмпирических и т.д. Иначе говоря, само системное движение представляет собой сложнейшую систему, иерархические связи между подсистемами которой, как, впрочем, и специфика ее многих подсистем, для нас пока еще во многом не ясны. Отсюда следует, во-первых, что отдельные системные подходы (по Берталанфи) действительно могут создаваться на основе не во всем системных и даже совсем не системных разработок и, во-вторых, что решение задачи четкого осознания различия и многообразия системных проблем, выделения основных сфер системных исследований становится в настоящее время важнейшим условием успешной разработки системного подхода.

В сжатом виде история развития системных идей представлена в табл. 1.1.

Основные вехи эволюции системных идей	**Основные положения**
Рождение понятия «система» (2500—2000 гг. до н. э.)	Слово «система» появилось в Древней Элладе и означало сочетание, организм, организация, союз. Выражало и некоторые акты деятельности (нечто, поставленное вместе, приведенное в порядок). Связано с формами социально-исторического бытия

Тезисы Демокрита (460—370 гг. до н. э.), Аристотеля (384—322 гг. до н. э)	Перенос значения слова с одного объекта на другой совершается поэтапно. Метафоризация (перенос скрытое уподобление, метафораобразное сближение слов на базе их переносного значения, например: «свинцовая туча») была начата греческим философом Демокритом. Он уподобил образование сложных тел из атомов с образованием слов из слогов. Аристотель трансформировал метафору в философской системе. Важно, что именно в античной философии был сформулирован тезис — целое больше суммы его частей (Философский словарь М.: Политиздат, 1980. С. 329)
Концепции эпохи Возрождения	Трактовка бытия как космоса сменяется на систему мира как независимое от человека, обладающее определенной организацией, иерархией, структурой Бытие становится не только предметом философского размышления (для постижения целостности), но и специально-научного анализа (каждая дисциплина вычленяет определенную область)
Идеи Н. Коперника (1473—1543)	Новая трактовка системности — в создании гелиоцентрической картины мира. Земля, как и другие планеты, обращается вокруг Солнца
Идеи Г. Галилея (1564—1642), И. Ньютона (1642—1727)	Галилей и Ньютон преодолели телеологизм (учение о конечных причинах) Николая Коперника в его астрономии, выработали определенную концептуальную систему с категориями — вещь и свойства, целое и часть... Вещь трактовалась как сумма отдельных свойств (забыли тезис античности???). Отношение выражало воздействие некоего предмета на другой, первый из которых являлся причиной, а второй — следствием. Очень важно: на первый план выдвигался каузальный, а не телеологический способ объяснения
Немецкая классическая	Глубокая и основательная разработка идеи системной организации научного знания. Структура научного знания

философия	стала предметом специального философского анализа
Идеи И. Ламберта (1728—1777)	Всякая наука, как и ее часть, предстает как система, трактуемая как целое!
Идеи И. Канта (1724—1804)	Кант не только осознал системный характер научного знания, но и превратил эту проблему в методологическую, выявив процедуры системного конструирования знания. Однако он считал, что принципы образования систем являются характеристиками лишь формы, а не содержания знания
Идеи И. Фихте (1762—1814)	Фихте поправил И. Канта, считая, что научное знание есть системное целое. Однако он ограничил системность знания систематичностью его формы. Это привело к отождествлению системности научного знания и его систематического изложения, т. е. внимание обращалось не на научное исследование, а на изложение знания
Идеи Г. Гегеля (1770—1831)	Гегель исходил из единства содержания и формы знания, тождества мысли и действительности. Трактовал становление системы в соответствии с принципом восхождения от абстрактного к конкретному. Но отождествляя метод и систему, телеологически истолковывая историю знания, он не смог предложить методологические средства для формирования системных образований
Теоретическое естествознание XIX—XX вв.	Различение объекта и предмета познания, повышение роли моделей в познании, фиксация наличия особых интегративных характеристик, исследование системообразующих принципов (порождение свойств целого из элементов и свойств элементов из целого), возможность предсказания!!!

Марксизм	Человек в процессе производства может действовать лишь так, как действует сама природа. Теоретики марксизма выдвинули принципы анализа системности научного знания: историзм, единство содержания и формы, трактовка системности как открытой системы
Идеи А. А. Богданова (1873—1928)	Богданов выразил многие важные идеи кибернетики, сформулированные Н. Виннером и У. Эшби, значительно раньше, хотя и в иной форме. Предвосхитил ОТС Л. Берталанфи в работе по тектологии (от гр. «строитель»). Основная идея — признание необходимости подхода к любому явлению со стороны его организованности (системности — других авторов). Под организованностью он понимает свойство целого быть больше суммы своих частей. Чем больше целое разнится от суммы, тем более оно организованно!!!
Идеи Л. Берталанфи (1901—1972)	Берталанфи первым из западных ученых разработал концепцию организма как открытой системы и сформулировал программу построения ОТС. Проводил мысль о неразрывности естественнонаучного [биологического) и философского (методологического) Сначала создал теорию открытых систем, граничащую с современной физикой, химией и биологией. Классическая термодинамика исследовала лишь закрытые системы. Организм представляет собой открытую систему, остающуюся постоянной при непрерывном изменении входящих в него веществ и энергии (так называемое состояние подвижного равновесия). Позже он обобщил идеи ТОС и выдвинул программу построения ОТС, являющейся всеобщей теорией организации. Проблемы организации, целостности, динамического взаимодействия были чужды классической физике. Он пришел к концепции синтеза наук, которую в противоположность «редукционизму», т.е. сведению всех наук к физике, он называет «перспективизмом». ОТС освобождает ученых от массового дублирования работ, экономя астрономические суммы денег и времени. Его недостатки: неполное определение «системы», отсутствие особенностей саморазвивающихся систем, теоретические исследования

	не всех видов «связи» и пр. Но главный недостаток: утверждение автора, что ОТС выполняет роль философии современной науки. Но это не так, ибо для философского учения с методах исследования необходимы совершение иные (новые) исходные понятия и иная направленность анализа: абстрактное и конкретное, специфически мысленное знание, связь знаний ОТС.
Концепции современности	Идеи СП нашли свое отражение в работах следующих авторов: Р. Акоффа, В. Афанасьева, С. Вира, И. Блауберга, Д. Бурчфилда, Д. Гвишиани, Г. Гуда, Д. Диксона, А. Зиновьева, Э. Квейда, В. Кинга, Д. Клиланда, В. Кузьмина, О. Ланге, В. Лекторского, В. Лефевра, Е. Липатова, Р. Макола, А. Малиновского, М. Месаровича, Б. Мильнера, Н. Овчинникова, С. Оптнера, Г. Поварова, Б. Радвига, А. Рапопорта, В. Розина, В. Садовского, М. Сетрова, В. Топорова, А. Уемова, Б. Флейшмана, Ч. Хитча, А. Холла, Б. Юдина, Ю. Черняка, Г. Щедровицкого, У. Эшби, Э. Юдина

Таблица 1.1 — История развития системных идей

Особый интерес представляет собой история развития системного подхода в технике.

Начиная с 20-х годов нашего века (и по сегодняшний день) появляются попытки построить социально-научные концепции в разных дисциплинах.

В биологии была создана организмические концепция, провозгласившая, что интегративные (целостные) характеристики не могут быть выведены из элементаризма, с крайней формой классического механистического атомизма. Здесь одним из главных тезисов системного подхода стал лозунг: в живом организме надо рассматривать не только множество связей, но и многообразие типов связей. Причинно-следственные связи перестали быть единственным видом связей, признаваемых наукой. Приобрели «права гражданства» функциональные, корреляционные, связи развития и др.

В психологии возникла новая концепция — гештальтпсихология, в основе которой лежит тезис: в психологических процессах важнейшую роль играют структурированные целые (гештальты).

В социологии можно выделить два основных подхода к исследованию общества. Это структурно-функциональный анализ, который исследует особенности развитого общества, определяющую роль способа производства по отношению к другим сторонам общественной жизни, противоречия между материальными и духовными явлениями жизни, специфические особенности и сложность выражения экономических отношений через взаимодействие политических, правовых, семейных, эмоциональных и других отношений, существующих в обществе.

Другой подход к исследованию социальных явлений — это генетический анализ. Его задачи — понимание общества как развивающегося целого, выделение качественных особенностей каждой ступени его развития. В конечном счете эти два способа исследования взаимно дополняют друг друга, позволяя понять общество как единое целое.

В технике выдвинуты общие проблемы синтеза многих различных факторов и подходов при конструировании сложных технических систем (ТС). Это проблемы «человек-машина», инженерной психологии, исследования операций и пр. Сама деятельность разработки ТС начинает выступать как сложная проблема, требующая специальных средств управления. Иными словами, развитие техники приводит к системной организованности самой деятельности, т.е. к требованию строгой взаимосвязи усилий и методов инженера и психолога, математика и врача, физика и экономиста.

Анализ исторического материала показывает, что стихийное становление системного подхода связано с техникой. В стихийном, неосознанном виде идея системности техники выражена уже в работах античных авторов, которые имели дело с относительно простыми механизмами. В качестве источника при рассмотрении этого периода в развитии техники используется трактат Марка Витрувия «Об архитектуре», который историки античности называют «энциклопедией техники античного периода». В описании конструкций механизмов у Витрувия достаточно полно раскрывается системный характер техники. Характеризуя функцию механизма, Витрувий далее рассматривает то, как связана функция объекта с тем определенным множеством взаимодействующих элементов, которое определяет эту функцию. Здесь Витрувий переходит уже к описанию структуры механизма. Причем важно отметить, что фиксируется не просто вообще взаимодействие элементов механизма, а упорядоченное расположение одних элементов относительно других.

В начале нашего века российский инженер П.К. Энгельмейер высказал мысль, которую, несколько перефразировав, можно передать так: в отношении к техническим изобретениям в интеллигентной публике замечается странность — принято восторгаться этими изобретениями, но в них не принято видеть деятельности, имеющей право быть поставленной рядом с деятельностью естествоиспытателя. В этом высказывании четко выражено исторически сложившееся отношение к технической деятельности. Очевидно, такое отношение возникает в связи тем, что в процессе развития познания центр внимания сосредоточивается на изучении естественнонаучной деятельности. Что же касается технической деятельности, то ученые, как правило, признавая важность тех или иных технических изобретений, не видели в этой деятельности предмета, достойного социального изучения. Этот факт нашел свое закрепление и в философии. Здесь уместно будет напомнить мысль В. И. Ленина, высказанную им в «Философских тетрадях»: «Продолжение дела Гегеля и Маркса должно состоять в диалектической обработке истории человеческой мысли, науки и техники» [Ленин В. И. // Полн. собр. соч. Т. 29. С. 131].

Итак, с эпохи античности продолжалось стихийное, неосознанное использование элементов системности, и то лишь в отдельных отраслях познания. Это составило первый этап исторического развития системного подхода.

Однако с середины XX века при появлении сложных и больших технических систем потребовалось специальное теоретическое обоснование методологического характера. Резко возросли комплексность и сложность проблем, некоторые из них стали глобальными (например, связь с помощью спутников). Усилилась зависимость между отдельными вопросами, которые раньше казались несвязанными. Актуальность решения проблем значительно возросла. Затраты на реализацию того или иного решения стали достигать многих десятков, сотен миллионов и даже миллиардов долларов, а риск неудачи становился все ощутимее. Потребовался учет все большего числа взаимосвязанных обстоятельств, а времени на решение становилось все меньше. Особенно это касалось разработки новой военной техники. Если раньше относительные затраты на вооружение были невелики, возможностей для выбора было мало, то фактически использовался принцип: «Ничего, кроме самого лучшего». Но с началом атомного века расходы на создание оружия возросли во много раз, и этот подход стал неприемлемым. Его постепенно заменял другой: «Только то, что необходимо, и за минимальную стоимость». Однако для реализации нового принципа нужно было уметь находить, оценивать и сравнивать альтернативы оружия. Потребовались методы, которые бы позволили анализировать сложные проблемы как целое, обеспечивали рассмотрение многих альтернатив, каждая из которых описывалась большим числом переменных, обеспечивали полноту каждой альтернативы, помогали вносить измеримость, давали возможность отражать объективные и субъективные неопределенности. Получившаяся в результате развития и обобщения широкая и универсальная методология решения проблем была названа ее авторами «системный анализ». Новая методология, созданная для решения военных проблем, была прежде всего использована в этой области.

Разработка и широкое применение системного анализа — заслуга знаменитой фирмы «РЭНД корпорейшн», основанной в 1947 г. Специалисты этой мощной корпорации выполнили ряд основополагающих исследований и разработок по СА, ориентированных на решение слабоструктурированных (смешанных) проблем Министерства обороны США. В 1948 г. Министерством ВВС была организована группа оценки систем оружия, а два года спустя — отдел анализа стоимости вооружения. Начавшееся в 1952 г. создание сверхзвукового бомбардировщика В-58 было первой разработкой, поставленной как система. Все это требовало выпуска монографической и учебной литературы. Первая книга по СА, не переведенная у нас, вышла и 1956 г. Ее издала РЭНД (авторы А. Кан и С. Манн). Через год появилась «Системотехника» Г. Гуда и Р. Макола » (издана у нас в 1962 г.), где изложена общая методика проектирования сложных технических систем. Методология СА была детально разработана и представлена в вышедшей в I960 г. книге Ч. Хитча и Р. Маккина «Военная экономика в ядерный век» (издана у нас в 1964 г.). В ней также приводится приложение к методам количественного сравнения альтернатив для решения проблем вооружения. В 1962 г. выходит один из самых лучших учебников по системотехнике (А. Холл «Опыт методологии для системотехники», переведенная у нас в 1975 г.), носящий не справочный или прикладной характер, а представляющий теоретическую разработку проблем системотехники. В 1965 г. появилась весьма обстоятельная книга Э. Квейда «Анализ сложных систем для решения военных проблем» (переведена в 1969 г.). В ней представлены основы новой научной дисциплины — анализа систем,— направленной на обоснование методов оптимального выбора при решении сложных проблем в условиях высокой неопределенности. Эта книга является переработанным изложением курса лекций по анализу систем, прочитанных работниками корпорации РЭНД для руководящих специалистов Министерства обороны и промышленности США. В 1965 г. вышла книга С. Оптнера «Системный анализ для решения деловых и промышленных проблем» (переведена в 1969 г.). Написанная лаконично, но насыщенная большим количеством новых идей, она дает полное и ясное представление о СА с характеристикой проблем делового мира, сущности систем и методологии решения проблем. Книга явилась одной из первых изданных у нас работ, освещающих состояние этой области в США.

Очень скоро выяснилось, что проблемы гражданские, проблемы фирм, маркетинга, аудита и прочие не только допускают, но и требуют обязательного применения этой методологии. Системный подход довольно быстро превратился в важный метод познания, в отличие от специальных приемов, характерных для разработки техники XVI—XIX вв. Это составило второй этап исторического развития системного подхода в технике.

Если при стихийном использовании системного подхода было главной целью изучение конечных результатов, то для второго этапа характерно переключение внимания на начальные стадии, связанные с выбором и обоснованием целей, их полезности, условий осуществления, связей с предыдущими процессами. Это потребовало знаний о структуре и функциях ТС, что повысило роль теоретических знаний. Если теоретическая деятельность первого этапа была направлена на описание и классификацию изучаемых объектов, то главными моментами второго этапа стали выявление механизмов функционирования ТС, а также знание условий, нарушающих их нормальную деятельность. Механизм функционирования включает исследование функций системы, определение связей функции со множеством взаимодействующих элементов, рассмотрение структуры ТС не как отношение (взаимосвязь, взаимодействие), а как определенным образом упорядоченное расположение одних элементов ТС относительно других (отношения между отношениями). Знание структуры и функций ТС является важным, но не достаточным условием для эффективного решения современных проблем. Надо обязательно соотнести цели субъекта с целями системы и выяснить, как скажется их реализация на функционировании ТС.

Современное развитие системного подхода идет в трех направлениях:

1. системологии как теории ТС;
2. системотехники как практики;
3. системного анализа как методологии.

Обобщенный материал по истории развития СП в технике представлен в табл. 1.2.

Сначала системный анализ базировался главным образом на применении сложных математических приемов. Спустя некоторое время ученые пришли к выводу, что математика неэффективна при анализе широких проблем со множеством неопределенностей, которые характерны дли исследования и разработки техники как единого целого. Поэтому стала вырабатываться концепция такого системного анализа, в котором упор делается преимущественно на разработку новых диалектических принципов научного мышления, логического анализа ТС с учетом их взаимосвязей и противоречивых тенденций. При таком подходе на первый план выдвигаются уже не математические методы, а сама логика системного анализа, упорядочение процедуры принятия решений. И видимо, не случайно, что в последнее время под системным подходом зачастую понимается некоторая совокупность системных принципов.

Предложенные варианты обще системных концепций строятся на различных предпосылках и отличаются разнообразием используемых средств. Именно факт выдвижения этих концепций превратил системный подход в научную реальность. И этому не препятствует отсутствие единой общепринятой теории систем.

История развития Исследований операций

Исследование операций (К) — это наука, которая занимается разработкой и практическим применением методов оптимального управления организационными системами. Предметом К являются системы организационного управления или организации, складаются из большого количества взаимодействующих между: собой подразделений, интересы которых не всегда согласуются между: собой и могут быть полностью или частично противоположными. К служит для количественного обоснования решений, принимаемых в организациях, и исходит из того, что качество решения можно количественно оценить с помощью одного или нескольких критериев качества. Как наука, К возникло перед Второй мировой войной, исходя из военных нужд, и в дальнейшем нашло широкое применение к решению практических задач в экономике и других отраслях. Характерными чертами операционного подхода являются: системность, комплексность, ориентация на принятие оптимального решения, телеологичнисть и компьютеризация.

Основными понятиями К являются: операция, оперируя сторона, активные средства операции, стратегии оперируя стороны, действующие факторы операции, состояние операции, оптимальное решение, критерий качества.

Модели К применяются для поиска оптимальных решений как в детерминированных, так и в стохастических системах.

По содержанию задачи К подразделяются на следующие: распределения ресурсов; транспортировки продуктов (выбора маршрутов); планирования и управления на сетях; формирование расписаний планирования и размещения; управление запасами, ремонта и замены оборудования; массового обслуживания; принятия решений в ситуациях с активным противодействием.

Исследование операций (К) — это теория математических моделей и методов получения оптимальных решений, направленная на обоснование целесообразности выбора той или

иной альтернативы из множества возможных в области целенаправленной деятельности человека.
После изучения темы студенты должны
знать: предмет и цель исследования операций, историю возникновения К; основные понятия, черты и последовательность реализации операционного подхода; основные типы задач ГО;
уметь: различать прямую и обратную задачи К; определять характер неопределенности в ситуациях принятия решений.

КЛЮЧЕВЫЕ ПОНЯТИЯ И СРОКИ
ў исследования операций
ў стратегии ОС
ў математическая модель операции
ў задачи распределения ресурсов
ў оптимальное решение
ў операция
ў задачи управления запасами
ў детерминированность
ў задачи массового обслуживания
ў действующие факторы операции
ў стохастичность
ў игровые задачи
1. Значение использования современных математических методов и моделей в управлении.
Развитие современного общества достиг того уровня, когда возникает настоятельная необходимость в разработке эффективных методов управления организационными системами различного назначения и разных уровней. Примерами таких систем являются отдельные производства, отрасли хозяйства, структуры управления (военные, государственные), хозяйственные комплексы и т. Д.
Объектом научной дисциплины «Исследование операций в экономике» является анализ функ-ния производственно-хозяйственных систем и разработка методов оптимального управления ими с использованием соответствующих математических моделей. Решение этих проблем достигается системным, всесторонним изучением процессов в исследуемых системах, синтезом качественных исследований и определенного математического аппарата в сочетании с широкими возможностями современных ЭВМ.
Под термином «операция» понимают любую совокупность мероприятий, направленных на исследование определенной цели. Исследование функционирования структуры или системы можно выполнять различными способами.
Историческая справка.
Развитие общественно-производственных отношений обусловил появление такой науки, как политическая экономия (мастерство управлять экономикой страны). Основателем классической политической экономии был У. Петти (1623—1687).
Первая в мире модель хозяйства в государстве была создана французским ученым Ф. Кене (1694—1774). Его знаменитая «Экономическая таблица» стала основой разработки многих математических моделей общественного производства.

Важную роль в развитии экономической науки XIX века сыграла так называемая математическая школа в политической экономии. Ее виднейшие представители -О. Курно, Г. Госсен, Я. Вальрас, В. Джеванс, Ф. Эджворт, В. Парето сделали значительный вклад в разработку проблем потребления, спроса и предложений, сбалансированности (равновесия) развития экономики. В дореволюционной России в конце XIX века были выполнены оригинальные экономико-математические исследования В. К. Дмитриева, В. И. Борткевича, РМ Орженецького и др. В. К. Дмитриев предложил модель полных хозяйственных затрат труда и сбалансированных цен. Весомый вклад в разработку экономико-математического моделирования сделал Е. Е. Слуцкий (1880—1948). В 1930—1950 рр. в СССР не было прогресса в развитии экономико-математических исследований через идеологические притеснения тоталитарного режима.
Развитие и исследования производственных связей и социальных отношений, практика хозяйствования требуют решения соответствующих задач управления системами. В 1930—1934 рр. талантливый русский ученый Н. Д. Кондратьев (1892—1938), находясь в политизоляторе, разработал динамическую модель макроэкономики, использовав аппарат дифференциальных уравнений. В 1938—1939 рр. ленинградский математик Л. В. Канторович (1912—1986), исследуя некоторые проблемы организации и планирования производства, сформулировал новый класс условно-экстремальных задач. Соответствующая отрасль прикладной математики впоследствии была названа «линейное программирование»; под этим термином понимают обоснование плана соответствующих экономических мер.
В 1960—1980 рр. исследования в области использования экономико-математических методов возродились и расширились. Вышли из печати фундаментальные труды: "Экономический расчет наилучшего использования ресурсов» Л. В. Канторовича (1959), "Измерение затрат и их результатов в социалистическом хозяйстве» В. В. Новожилова (1959), «Экономико-математические методы и модели» В . С. Немчин-ва (1962).
В 70-е годы начались работы по внедрению достижений экономико-математического моделирования и теории исследования операций в разработку автоматизированных систем планирования и управления. Основное внимание исследователей в странах СНГ сейчас сосредоточено на решении проблем переходного периода. Одной из важных проблем западноевропейских и американских ученых является исследование социально-экономических вопросов развития рыночных отношений на планетарном уровне.
Изучение основных положений теории исследования операций в экономике будет способствовать формированию конструктивного мировоззрения у каждого будущего работника любого уровня социальной иерархии: воплощение любой якогозаходу начинается с плана его реализации, требует времени и средств; необходимо не только работать, но и оценивать эффективность своей.

2. Основные принципы системного анализа и теории принятия решений

1.1. Общие положения

Человек наделён сознанием, существо свободное и обречено на выбор решений, стараясь сделать всё наилучшим образом. В наиболее общем смысле теория принятия оптимальных решений представляет собой совокупность математических и численных методов, ориентированных на нахождение наилучших вариантов из множества альтернатив и позволяющих избежать их полного перебора. Ввиду того, что размерность практических задач, как правило, достаточно велика, а расчеты в соответствии с алгоритмами оптимизации требуют значительных затрат времени, то методы принятия оптимальных решений главным образом ориентированы на реализацию их с помощью ЭВМ. (Примечание редактора сайта Б.Майорова: не только ЭВМ, первый вариант решения можно оценить на пальцах).

Практическая потребность общества в научных основах принятия решений возникла с развитием науки и техники только в XVIII веке Началом науки "Теория принятия решений" следует считать работу Жозефа Луи Лагранжа, смысл которой заключался в следующем: сколько земли должен брать на лопату землекоп, чтобы его сменная производительность была наибольшей. Оказалось, что утверждение "бери больше, кидай дальше" неверен. Бурный рост технического прогресса, особенно во время и после второй мировой войны, ставил все новые и новые задачи, для решения которых привлекались и разрабатывались новые научные методы. Можно выделить следующие научно-технические предпосылки становления "Теории принятия решений":

- *удорожание "цены ошибки"*. Чем сложнее, дороже, масштабнее планируемое мероприятие, тем менее допустимы в нем "волевые" решения и тем важнее становятся научные методы, позволяющие заранее оценить последствия каждого решения, заранее исключить недопустимые варианты и рекомендовать наиболее удачные;
- *ускорение научно-технической революции техники и технологии*. Жизненный цикл технического изделия сократился настолько, что "опыт" не успевал накапливаться и требовалось применение более развитого математического аппарата в проектировании;
- *развитие ЭВМ.* Размерность и сложность реальных инженерных задач не позволяло использовать аналитические метода.

Как часто это бывает, эта наука, с одной стороны, стала определенной ветвью других более общих наук (теория систем, системный анализ, кибернетика и т.д.), а с другой, стала синтезом определенных фундаментальных более частных наук (исследование операций, оптимизация и т.д.), создав при этом и собственную методологию.

Инженерное дело теснейшим образом связано с совокупностями объектов, которые принято называть *сложными системами*, которые характеризуются многочисленными и разнообразными по типу связями между отдельно существующими элементами системы и наличием у системы функции назначения, которой нет у составляющих ее частей. На первый взгляд каждая сложная система имеет уникальную организацию. Однако более детальное изучение способно выделить общее в системе команд ЭВМ, в процессах проектирования лесной машины, самолета и космического корабля.

В научно-технической литературе существует ряд термином, имеющих отношение к исследованию сложных систем.

Наиболее общий термин "*теория систем*" относится ко всевозможным аспектам исследования систем. Ее основными частями являются

- *системный анализ*, который понимается как исследование проблемы принятия решения в сложной системе,

- *кибернетика*, которая рассматривается как наука об управлении и преобразовании информации.

Здесь следует заметить, что понятие *управления* не совпадает с *принятием решения*. Условная граница между кибернетикой и системным анализом состоит в том, что первая изучает отдельные и строго формализованные процессы, а системный анализ - совокупность процессов и процедур.

Очень близкое к термину "системный анализ" понятие - "*исследование операций*", которое традиционно обозначает математическую дисциплину, охватывающую исследование математических моделей для выбора величин, оптимизирующих заданную математическую конструкцию (критерий). Системный анализ может сводиться к решению ряда задач исследования операций, но обладает свойствами, не охватываемыми этой дисциплиной. Однако в зарубежной литературе термин "исследование операций" не является чисто математическим и приближается к термину "системный анализ". Широкая опора системного анализа на исследование операций приводит к таким его математизированным разделам, как

- постановка задач принятия решения;
- описание множества альтернатив;
- исследование многокритериальных задач;
- методы решения задач оптимизации;
- обработка экспертных оценок;
- работа с макромоделями системы.

1.2. Основные понятия системного анализа

Системный анализ - наука, занимающаяся проблемой принятия решения в условиях анализа большого количества информации различной природы.

Из определения следует, что целью применения системного анализа к конкретной проблеме является повышение степени обоснованности принимаемого решения, расширение множества вариантов, среди которых производится выбор, с одновременным указанием способов отбрасывания заведомо уступающим другим.

В системном анализе выделяют

- методологию;
- аппаратную реализацию;
- практические приложения.

Методология включает **определения** используемых понятий и **принципы системного подхода**.

Дадим основные **определения** системного анализа.

Элемент - некоторый объект (материальный, энергетический, информационный), который обладает рядом важных для нас свойств, но внутреннее строение (содержание) которого безотносительно к цели рассмотрения.

Связь - важный для целей рассмотрения обмен между элементами веществом, энергией, информацией.

Система - совокупность элементов, которая обладает следующими признаками:

- связями, которые позволяют посредством переходов по ним от элемента к элементу соединить два любых элемента совокупности;
- свойством, отличным от свойств отдельных элементов совокупности.

Практически любой объект с определенной точки зрения может быть рассмотрен как система. Вопрос состоит в том, насколько целесообразна такая точка зрения.

Большая система - система, которая включает значительное число однотипных элементов и однотипных связей. В качестве примера можно привести трубопровод. Элементами последнего будут участки между швами или опорами. Для расчетов на прочность по методу конечных элементов элементами системы считаются небольшие участки трубы, а связь имеет силовой (энергетический) характер - каждый элемент действует на соседние.

Сложная система - система, которая состоит из элементов разных типов и обладает разнородными связями между ними. В качестве примера можно привести ЭВМ, лесной трактор или судно.

Автоматизированная система - сложная система с определяющей ролью элементов двух типов:

- в виде технических средств;
- в виде действия человека.

Для сложной системы автоматизированный режим считается более предпочтительным, чем автоматический. Например, посадка самолета или захват дерева харвестерной головкой выполняется при участии человека, а автопилот или бортовой компьютер используется лишь на относительно простых операциях. Типична также ситуация, когда решение, выработанное техническими средствами, утверждается к исполнению человеком.

Структура системы - расчленение системы на группы элементов с указанием связей между ними, неизменное на все время рассмотрения и дающее представление о системе в целом. Указанное расчленение может иметь материальную, функциональную, алгоритмическую или другую основу. Пример материальной структуры - структурная схема сборного моста, которая состоит из отдельных, собираемых на месте секций и указывает только эти секции и порядок их соединения. Пример функциональной структуры - деление двигателя внутреннего сгорания на системы питания, смазки, охлаждения, передачи крутящего момента. Пример алгоритмической структуры - алгоритм программного средства, указывающего последовательность действий или инструкция, которая определяет действия при отыскании неисправности технического устройства.

Структура системы может быть охарактеризована по имеющимся в ней типам связей. Простейшими из них являются последовательное, параллельное соединение и обратная связь (рис.1.1).

Декомпозиция - деление системы на части, удобное для каких-либо операций с этой системой. Примерами будут: разделение объекта на отдельно проектируемые части, зоны обслуживания; рассмотрение физического явления или математическое описание отдельно для данной части системы.

Иерархия - структура с наличием подчиненности, т.е. неравноправных связей между элементами, когда воздействие в одном из направлений оказывают гораздо большее влияние на элемент, чем в другом. Виды иерархических структур разнообразны, но важных для практики иерархических структур всего две - древовидная и ромбовидная (рис.1.2).

Древовидная структура наиболее проста для анализа и реализации. Кроме того, в ней всегда удобно выделять иерархические уровни - группы элементов, находящиеся на одинаковом удалении от верхнего элемента. Пример древовидной структуры - задача проектирования технического объекта от его основных характеристик (верхний уровень) через проектирование основных частей, функциональных систем, групп агрегатов, механизмов до уровня отдельных деталей.

Принципы системного подхода - это положения общего характера, являющиеся обобщением опыта работы человека со сложными системами. Их часто считают ядром методологии. Известно около двух десятков таких принципов, ряд из которых целесообразно рассмотреть:

- принцип конечной цели: абсолютный приоритет конечной цели;
- принцип единства: совместное рассмотрение системы как целого и как совокупности элементов;
- принцип связности: рассмотрение любой части совместно с ее связями с окружением;
- принцип модульного построения: полезно выделение модулей в системе и рассмотрение ее как совокупности модулей;
- принцип иерархии: полезно введение иерархии элементов и(или) их ранжирование;
- принцип функциональности: совместное рассмотрение структуры и функции с приоритетом функции над структурой;
- принцип развития: учет изменяемости системы, ее способности к развитию, расширению, замене частей, накапливанию информации;
- принцип децентрализации: сочетание в принимаемых решениях и управлении централизации и децентрализации;
- принцип неопределенности: учет неопределенностей и случайностей в системе.

Аппаратная реализация включает стандартные приемы моделирования принятия решения в сложной системе и общие способы работы с этими моделями. Модель строится в виде связных множеств отдельных процедур. Системный анализ исследует как организацию таких множеств, так и вид отдельных процедур, которые максимально приспосабливают для принятия согласующихся и управленческих решений в сложной системе.

Модель принятия решения чаще всего изображается в виде схемы с ячейками, связями между ячейками и логическими переходами. Ячейки содержат конкретные действия - процедуры. Совместное изучение процедур и их организации вытекает из того, что без учета содержания и особенностей ячеек создание схем оказывается невозможным. Эти схемы определяют стратегию принятия решения в сложной системе. Именно с проработки связанного множества основных процедур принято начинать решение конкретной прикладной задачи.

Отдельные же процедуры (операции) принято классифицировать на *формализуемые* и *неформализуемые*. В отличие от большинства научных дисциплин, стремящихся к формализации, системный анализ допускает, что в определенных ситуациях неформализуемые решения, принимаемые человеком, являются более предпочтительными. Следовательно, системный анализ рассматривает в совокупности формализуемые и неформализуемые процедуры, и одной из его задач является определение их оптимального соотношения.

Формализуемые стороны отдельных операций лежат в области прикладной математики и использования ЭВМ. В ряде случаев математическими методами исследуется связное множество процедур и производится само моделирование принятие решения. Все это позволяет говорить о математической основе системного анализа. Такие области прикладной математики, как исследование операций и системное программирование, наиболее близки к системной постановке вопросов.

Практическое приложение системного анализа чрезвычайно обширно по содержанию. Важнейшими разделами являются научно-технические разработки и различные задачи экономики. Ссылки на системность исследований, анализа, подхода включает биологию, экологию, военное дело, психологию, социологию, медицину, управление государством и регионом, лесное и сельское хозяйство, обучение и многое другое.

1.3. Основные понятия исследования операций

Операцией называется всякое мероприятие (система действий), объединенное единым замыслом и направленное к достижению какой-то цели.

Цель исследования операций - предварительное количественное обоснование оптимальных решений.

Всякий определенный выбор зависящих от нас параметров называется *решением*. *Оптимальным* называются *решения*, по тем или другим признакам предпочтительные перед другими.

Параметры, совокупность которых образует решение, называются *элементами решения*.

Множеством допустимых решений называются заданные условия, которые фиксированы и не могут быть нарушены.

Показатель эффективности - количественная мера, позволяющая сравнивать разные решения по эффективности.

Все решения принимаются всегда на основе информации, которой располагает *лицо принимающее решение* (ЛПР).

Каждая задача в своей постановке должна отражать структуру и динамику знаний ЛПР о множестве допустимых решений и о показателе эффективности.

Задача называется *статической*, если принятие решения происходит в наперед известном и не изменяющемся информационном состоянии. Если информационное состояние в ходе принятия решения сменяют друг друга, то задача называется *динамической*.

Информационные состояния ЛПР могут по-разному характеризовать его физическое состояние:

- Если информационное состояние состоит из единственного физического состояния, то задача называется *определенной*.
- Если информационное состояние содержит несколько физических состояний и ЛПР кроме их множества знает еще и вероятности каждого из этих физических состояний, то задача называется *стохастической* (частично неопределенной).
- Если информационное состояние содержит несколько физических состояний, но ЛПР кроме их множества ничего не знает о вероятности каждого из этих физических состояний, то задача называется *неопределенной*.

1.4. Постановка задач принятия оптимальных решений

Несмотря на то, что методы принятия решений отличаются универсальностью, их успешное применение в значительной мере зависит от профессиональной подготовки специалиста, который должен иметь четкое представление о специфических особенностях изучаемой системы и уметь корректно поставить задачу. Искусство постановки задач постигается на примерах успешно реализованных разработок и основывается на четком представлении преимуществ, недостатков и специфики различных методов оптимизации. В первом приближении можно сформулировать следующую последовательность действий, которые составляют содержание процесса постановки задачи:

- *установление границы подлежащей оптимизации системы,* т.е. представление системы в виде некоторой изолированной части реального мира. Расширение границ системы повышает размерность и сложность многокомпонентной системы и, тем самым, затрудняет ее анализ. Следовательно, в инженерной практике следует к декомпозиции сложных систем на подсистемы, которые можно изучать по отдельности без излишнего упрощения реальной ситуации;
- *определение показателя эффективности*, на основе которого можно оценить характеристики системы или ее проекта с тем, чтобы выявить "наилучший" проект или множество "наилучших" условий функционирования системы. В инженерных приложениях обычно выбираются показатели экономического (издержки, прибыль и т.д.) или технологического (производительность, энергоемкость, материалоемкость и т.д.) характера. "Наилучшему" варианту всегда соответствует экстремальное значение показателя эффективности функционирования системы;
- *выбор внутрисистемных независимых переменных,* которые должны адекватно описывать допустимые проекты или условия функционирования системы и способствовать тому, чтобы все важнейшие технико-экономические решения нашли отражение в формулировке задачи;
- *построение модели,* которая описывает взаимосвязи между переменными задачи и отражает влияние независимых переменных на значение показателя эффективности. В самом общем случае структура модели включает основные

уравнения материальных и энергетических балансов, соотношения, связанные с проектными решениями, уравнения, описывающие физические процессы, протекающие в системе, неравенства, которые определяют область допустимых значений независимых переменных и устанавливают лимиты имеющихся ресурсов. Элементы модели содержат всю информацию, которая обычно используется при расчете проекта или прогнозировании характеристик инженерной системы. Очевидно, процесс построения модели является весьма трудоемким и требует четкого понимания специфических особенностей рассматриваемой системы.

Несмотря на то, модели принятия оптимальных решений отличаются универсальностью, их успешное применение зависит от профессиональной подготовки инженера, который должен иметь полное представление о специфике изучаемой системы. Основная цель рассмотрения приводимых ниже примеров - продемонстрировать разнообразие постановок оптимизационных задач на основе общности их формы.

Все оптимизационные задачи имеют общую структуру. Их можно классифицировать как задачи минимизации(максимизации) M-векторного векторного показателя эффективности $W_m(x)$, $m=1,2,...,M$, N-мерного векторного аргумента $x=(x_1,x_2,...,x_N)$, компоненты которого удовлетворяют системе ограничений-равенств $h_k(x)=0$, $k=1,2...K$, ограничений-неравенств $g_j(x)>0$, $j=1,2,...J$, областным ограничениям $x_{li}<x_i<x_{ui}$, $i=1,2...N$.

Все задачи принятия оптимальных решений можно классифицировать в соответствии с видом функций и размерностью $W_m(x)$, $h_k(x)$, $g_j(x)$ и размерностью и содержанием вектора x:

- *одноцелевое принятие решений* - $W_m(x)$ - скаляр;
- *многоцелевое принятие решений* - $W_m(x)$ - вектор;
- *принятие решений в условиях определенности* - исходные данные - детерминированные;
- *принятие решений в условиях неопределенности* - исходные данные - случайные.

Наиболее разработан и широко используется на практике аппарат одноцелевого принятия решений в условиях определенности, который получил название *математического программирования.* Более подробно задачи линейного программирования ($W(x)$, $h_k(x)$, $g_j(x)$ - линейны) изложены в главе 2, нелинейного программирования ($W(x)$, $h_k(x)$, $g_j(x)$ - нелинейны) - в главе 3, целочисленного программирования (x - целочисленны) - в главе 4, динамического программирования (x - зависят от временного фактора) - в главе 5.

Математический аппарат одноцелевого принятия решений в условиях неопределенности, изложенный в главе 6, представляет собой стохастическое программирование (известны законы распределения случайных величин), теории игр и статистических решений (закон распределения случайных величин неизвестен).

Методы принятия многоцелевых решений изложены в 7 главе.

Рассмотрим процесс принятия решений с самых общих позиций. Психологами установлено, что *решение* не является начальным процессом творческой деятельности. Оказывается, непосредственно акту решения предшествует тонкий и обширный процесс работы мозга, который формирует и предопределяет направленность решения. В этот этап, который можно назвать "предрешением" входят следующие элементы:

- *мотивация*, то есть желание или необходимость что-то сделать. Мотивация определяет цель какого-либо действия, используя весь прошлый опыт, включая результаты;
- возможность неоднозначности результатов;
- возможность неоднозначности способов достижения результатов, то есть свобода выбора.

После этого предварительного этапа следует, собственно, этап *принятия решения.* Но на нем процесс не заканчивается, т.к. обычно после принятия решения следует оценка результатов и корректировка действий. Таким образом, принятие решений следует воспринимать не как единовременный акт, а как последовательный процесс.

Выдвинутые выше положения носят достаточно общий характер, обычно подробно исследуемый психологами. Более близкой с точки зрения инженера будет следующая схема процесса принятия решения. Эта схема включает в себя следующие компоненты:

- анализ исходной ситуации;
- анализ возможностей выбора;
- выбор решения;
- оценка последствий решения и его корректировка.

1.5. Принятие решений в условиях неопределенности

Постановка задачи

В изложенных выше материалах речь шла о постановках и методах решения задач, не содержащих неопределенностей. Однако, как правило, большинство реальных инженерных задач содержит в том или ином виде неопределенность. Можно даже утверждать, что решение задач с учетом разного вида неопределенностей является *общим случаем*, а принятие решений без их учета - *частным*. Однако, из-за концептуальных и методических трудностей в настоящее время не существует единого методологического подхода к решению таких задач. Тем не менее, накоплено достаточно большое число методов формализации постановки и принятия решений с учетом неопределенностей. При использовании этих методов следует иметь в виду, что все они носят рекомендательный характер и выбор окончательного решения всегда остается за человеком (ЛПР).

Как уже указывалось, при решении конкретных задач с учетом неопределенностей инженер сталкивается с разными их типами. В исследовании операций принято различать три типа неопределенностей:

- неопределенность целей;
- неопределенность наших знаний об окружающей обстановке и действующих в данном явлении факторах (неопределенность природы);
- неопределенность действий активного или пассивного партнера или противника.

В приведенной выше классификации тип неопределенностей рассматривается с позиций того или иного элемента математической модели. Так, например, неопределенность целей отражается при постановке задачи на выборе либо отдельных критериев, либо всего вектора полезного эффекта.

С другой стороны, два другие типа неопределенностей влияют, в основном, на составление целевой функции уравнений ограничений и метода принятия решения. Конечно, приведенное выше утверждение является достаточно условным, как, впрочем, и любая классификация. Мы приводим его лишь с целью выделить еще некоторые особенности неопределенностей, которые надо иметь в виду в процессе принятия решений.

Дело в том, что кроме рассмотренной выше классификации неопределенностей надо учитывать их тип (или "род") с точки зрения отношения к случайности.

По этому признаку можно различать стохастическую (вероятностную) неопределенность, когда неизвестные факторы статистически устойчивы и поэтому представляют собой обычные объекты теории вероятностей - случайные величины (или случайные функции, события и т.д.). При этом должны быть известны или определены при постановке задачи все необходимые статистический характеристики (законы распределения и их параметры).

Примером таких задач могут быть, в частности, система технического обслуживания и ремонта любого вида техники, система организации рубок ухода и т.д.

Другим крайним случаем может быть неопределенность нестохастического вида (по выражению Е.С.Вентцель [15] - "дурная неопределенность"), при которой никаких предположений о стохастической устойчивости не существует. Наконец, можно говорить о промежуточном типе неопределенности, когда решение принимается на основании каких-либо гипотез о законах распределения случайных величин. При этом ЛПР должен иметь в виду опасность несовпадения его результатов с реальными условиями. Эта опасность несовпадения формализуется с помощью коэффициентов риска.

Принятие решений в условиях риска

Как указывалось выше, с точки зрения знаний об исходных данных в процессе принятия решений можно представить два крайних случая: определенность и неопределенность. В некоторых случаях неопределенность знаний является как бы "неполной" и дополняется некоторыми сведениями о действующих факторах, в частности, знанием законов распределения описывающих их случайных величин. Этот промежуточный случай соответствует ситуации *риска*. Принятие решений в условиях риска может быть основано на одном из следующих критериев:

- критерий ожидаемого значения;
- комбинации ожидаемого значения и дисперсии;
- известного предельного уровня;
- наиболее вероятного события в будущем.

Рассмотрим более подробно применение этих критериев.

Критерий ожидаемого значения (КОЗ).

Использование КОЗ предполагает принятие решения, обуславливающего максимальную прибыль при имеющихся исходных данных о вероятности полученного результата при том или другом решении. По существу, КОЗ представляет собой выборочные средние значения случайной величины. Естественно, что достоверность получаемого решения при этом будет зависеть от объема выборки. Так, если обозначить

КОЗ - $E(x_1, x_2,..., x_n)$, (6.1)

где

$x_1, x_2,..., x_n$ - принимаемые решения при их количестве, равном n, то

$E(x_i)$ □□□ $M(x_i)$, (6.2)

где

$M(x_i)$ - математическое ожидание критерия.

Таким образом, КОЗ может применяться, когда однотипные решения в сходных ситуациях приходится принимать большое число раз.

Приведем пример использования этого критерия для принятия решения.

Критерий "ожидаемого значения - дисперсия".

Как указывалось выше, КОЗ имеет область применения, ограниченную значительным числом однотипных решений, принимаемых в аналогичных ситуациях. Этот недостаток можно устранить, если применять комбинацию КОЗ и выборочной дисперсии s^2. Возможным критерием при этом является минимум выражения

E(Z, □) = E(Z) □ k□ U(z), (6.5)

где

E(Z, □) - критерий "ожидаемого значения - дисперсия";

k - постоянный коэффициент;

$U(Z) = m_Z/S$ - выборочный коэффициент вариации;

m_Z - оценка математического ожидания;

S - оценка среднего квадратического ожидания.

Знак "минус" ставится в случае оценки прибыли, знак "плюс" - в случае затрат.

Из зависимости (6.5) видно, что в данном случае точность предсказания результата повышается за счет учета возможного разброса значений E(Z), то есть введения своеобразной "страховки". При этом степень учета этой страховки регулируется коэффициентом k, который как бы управляет степенью учета возможных отклонений. Так, например, если для ЛПР имеет большое значение ожидаемые потери прибыли, то k>>1 и при этом существенно увеличивается роль отклонений от ожидаемого значения прибыли E(Z) за счет дисперсии.

Критерий предельного уровня.

Этот критерий не имеет четко выраженной математической формулировки и основан в значительной степени на интуиции и опыте ЛПР. При этом ЛПР на основании субъективных соображений определяет наиболее приемлемый способ действий. Критерий предельного уровня обычно не используется, когда нет полного представления о множестве возможных альтернатив. Учет ситуации риска при этом может производиться за счет введения законов распределений случайных факторов для известных альтернатив.

Несмотря на отсутствие формализации критерием предельного уровня пользуются довольно часто, задаваясь их значениями на основании экспертных или опытных данных.

Критерий наиболее вероятного исхода.

Этот критерий предполагает замену случайной ситуации детерминированной путем замены случайной величины прибыли (или затрат) единственным значением, имеющим *наибольшую вероятность реализации*. Использование данного критерия, также как и в предыдущем случае в значительной степени опирается на опыт и интуицию. При этом необходимо учитывать два обстоятельства, затрудняющие применение этого критерия:

- критерий нельзя использовать, если *наибольшая вероятность* события недопустимо мала;
- применение критерия невозможно, если несколько значений вероятностей возможного исхода равны между собой.

Учет неопределенных факторов, заданных законом распределения.

Случай, когда неопределенные факторы заданы распределением, соответствует ситуации риска. Этот случай может учитываться двумя путями. Первый - анализом адаптивных возможностей, позволяющих реагировать на конкретные исходы; второй - методически, при сопоставлении эффективности технических решений. Суть первого подхода заключается в том, что законы распределения отдельных параметров на этапе проектирования могут быть определены с достаточной степенью приближения на основе сопоставления с аналогами, из физических соображений или на базе статистических данных и данных прогнозов.

Методический учет случайных факторов, заданных распределением, может быть выполнен двумя приемами: заменой случайных параметров их математическими ожиданиями (сведением стохастической задачи к детерминированной) и "взвешиванием" показателя качества по вероятности (этот прием иногда называют "оптимизация в среднем").

Первый прием предусматривает определение математического ожидания случайной величины v - M(v) и определение зависимости W(M(v)), которая в дальнейшем оптимизируется по u. Однако сведение к детерминированной схеме может быть осуществлено в тех случаях, когда диапазон изменения параметра u невелик или когда зависимость W(u) линейна или близка к ней.

Второй прием предусматривает определение W в соответствии с зависимостями соответственно для дискретных и непрерывных величин:

$$W = \sum_{i=1} W(u_i)P(u_i) \quad ; (6.6)$$

$W = \int W(u) f(u) du$, (6.7)

где

$P(u_i)$ - ряд распределений случайной величины u_i;

$f(u_i)$ - плотность распределения случайной величины u.

При описании дискретных случайных величин наиболее часто используют распределения Пуассона, биноминальное. Для непрерывных величин основными распределениями являются нормальное, равномерное и экспоненциальное.

Постановка задачи стохастического программирования

При перспективном и оперативном планировании работы предприятия возникает необходимость в учете ряда случайных факторов, существенно влияющих на процесс производства. К таким факторам относятся спрос, который не всегда может быть предсказуем, непредусмотренные сбои в поступлении сырья, энергии, рабочей силы, неисправности и аварии оборудования. Еще больше случайных факторов необходимо учитывать при планировании производства, эффективность которого зависит от климатических условий, урожайности и т.д. Поэтому, например, задачи планирования лесного производства целесообразно ставить и исследовать в терминах и понятиях стохастического программирования, когда элементы задачи линейного программирования (матрица коэффициентов A, вектора ресурсов b, вектора оценок c) часто оказываются случайными. Подобного типа задачи ЛП принято классифицировать как задачи стохастического программирования (СП).

Подходы к постановке и анализу стохастических задач существенно различаются в зависимости от последовательности получения информации - в один прием или по частям. При построении стохастической модели важно также знать, необходимо ли принять единственное решение, не подлежащее корректировке, или можно по мере накопления информации один или несколько раз корректировать решение. В соответствии с этим в стохастическом программировании исследуются одноэтапные, двухэтапные и многоэтапные задачи.

В *одноэтапных* задачах решение принимается один раз и не корректируется. Они различаются по показателям качества решения (по целевым функциям), по характеру ограничений и по виду решения.

Задача СП может быть сформулирована в М- и Р- постановках по отношению к записи целевой функции и ограничений.

Случайны элементы вектора c (целевая функция).

При М-постановке целевая функция W записывается в виде

$$W = M\left(\sum_{j=1}^{n} c_j x_j\right) \to \min(\max)$$, (6.8)

что означает оптимизацию математического ожидания целевой функции. От математического ожидания целевой функции можно перейти к математическому ожиданию случайной величины c_j

$$W = M(\sum_{j=1}^{n} c_j x_j) = \sum_{j=1}^{n} \bar{c}_j x_j \to \min(\max) \quad (6.9)$$

При Р- постановке имеем:

- при максимизации

$$W = P(\sum_{j=1}^{n} c_j x_j \geq W_{min}) \to \max, \quad (6.10)$$

где

W_{min} - предварительно заданное допустимое наихудшее (минимальное) значение целевой функции.

- при минимизации

$$W = P(\sum_{j=1}^{n} c_j x_j \leq W_{max}) \to \max, \quad (6.11)$$

где

W_{max} - предварительно заданное допустимое наихудшее (максимальное) значение целевой функции.

Суть Р-постановки заключается в том, что необходимо найти такие значения x_j, при которых максимизируется вероятность того, что целевая функция будет не хуже предельно допустимого значения.

Ограничения задачи, которые должны выполняться при всех реализациях параметров условий задачи, называются *жесткими* ограничениями. Часто возникают ситуации, в которых постановка задачи позволяет заменить жесткие ограничения их усреднением по распределению случайных параметров. Такие ограничения называют *статистическими*:

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j \leq \bar{b}_i . \quad (6.12)$$

В тех случаях, когда по содержательным соображениям можно допустить, чтобы невязки в условиях не превышали заданных с вероятностями, небольшими $\alpha_i>0$, говорят о стохастических задачах с *вероятностными* ограничениями:

$$P\left\{\sum_{j=1}^{n} \bar{a}_{ij} x_j \leq \bar{b}_i\right\} \geq \alpha_i, \quad (6.13)$$

т.е. вероятность выполнения каждого заданного ограничения должна быть не менее назначенной величины α_i. Параметры α_i предполагаются заданными или являются решениями задачи более высокого уровня.

Представленные задачи как в М-, так и в Р- постановках непосредственно решены быть не могут. Возможным методом решения этих задач является переход к их детерминированным эквивалентам. В основе этого перехода лежит использование закона распределения случайной величины. В инженерной практике наиболее часто используется нормальный закон распределения, поэтому дальнейшие зависимости приведем для этого случая.

Принимаем, что a_{ij}, b_i, c_j подчинены нормальному закону распределения. В этом случае будет справедлива следующие детерминированные постановки:

- Р - постановка целевой функции, максимизация:

$$W = \frac{\sum_{j=1}^{n} \bar{c}_j x_j - W_{min}}{\sqrt{\sum_{j=1}^{n} \sigma_j^2 x_j^2}} \rightarrow max, \qquad (6.14)$$

где

$\bar{c}_j$ и σ_j - математическое ожидание и среднее квадратическое отклонение случайной величины c_j.

- Р - постановка целевой функции, минимизация:

$$W = \frac{W_{max} - \sum_{j=1}^{n} \bar{c}_j x_j}{\sqrt{\sum_{j=1}^{n} \sigma_j^2 x_j^2}} \rightarrow max. \qquad (6.15)$$

- *Вероятностные* ограничения:

$$\sum_{j=1}^{n} \bar{a}_{ij} x_j \le \bar{b}_i - t_{\alpha_i} \sqrt{\sum_{j=1}^{n} \sigma_{ij}^2 x_j^2 + \theta_i^2},$$

где

$\bar{a}_{ij}, \sigma_{ij}^2, \bar{b}_i, \theta_i^2$ - соответственно, математические ожидания и дисперсии случайных величин a_{ij} и b_i;

t_{α_i} - значение центрированной нормированной случайной величины в нормальном законе распределения, соответствующей заданному уровню вероятности соблюдения ограничений α_i.

Сделаем несколько замечаний к приведенным зависимостям:

- задача стохастического программирования сведена к задаче нелинейной оптимизации и может быть решена одним из рассматриваемых ранее методов;
- сравнение ограничения ресурса в стохастическом программировании и аналогичным ограничением в задаче линейного программирования показывает, что учет случайного характера величин a_{ij} и b_i приводит к уменьшению располагаемого ресурса на величину

$$t_{\alpha}\sqrt{\sum_{j=1}^{n}\sigma_{ij}^{2}x_{j}^{2}+\sigma_{i}^{2}}, \quad (6.16)$$

т.е. к необходимости в дополнительном ресурсе. Однако этот дополнительный ресурс может оказаться неиспользованным, но для гарантированного выполнения плана его иметь необходимо.

Метод статистического моделирования

Приведенные формулы (6.6) и (6.7) могут быть использованы для систем независимых случайных величин. Однако для технических систем, как правило, случайные параметры являются зависимыми. Причем эта зависимость не функциональная, а корреляционная. Поэтому для анализа случайных факторов, заданных распределением, широкое применение нашли теория марковских процессов и метод статистического моделирования (метод Монте-Карло).

В задачах принятия оптимальных решений широкое применение получил метод Монте-Карло. Основными особенностями этого метода, основанного на многократном повторении одного и того же алгоритма для каждой случайной реализации, являются: универсальность (метод не накладывает практически никаких ограничений на исследуемые параметры, на вид законов распределения); простота расчетного алгоритма; необходимость большого числа реализаций для достижения хорошей точности; возможность реализации на его основе процедуры поиска оптимальных параметров проектирования. Отметим основные факторы, определившие применение метода статистического моделирования в задачах исследования качества при проектировании: метод применим для задач, формализация которых другими методами затруднена или даже невозможна; возможно применение этого метода для машинного эксперимента над не созданной в натуре системы, когда натурный эксперимент затруднен, требует больших затрат времени и средств или вообще не допустим по другим соображениям.

Учет неопределенных пассивных условий

Неопределенные факторы, закон распределения которых неизвестен, являются наиболее характерными при исследовании качества адаптивных систем. Именно на этот случай следует ориентироваться при выборе гибких конструкторских решений. Методический учет таких факторов базируется на формировании специальных критериев, на основе которых принимаются решения. Критерии Вальда, Сэвиджа, Гурвица и Лапласа уже давно и прочно вошли в теорию принятия решений.

В соответствии с *критерием Вальда* в качестве оптимальной выбирается стратегия, гарантирующая выигрыш не меньший, чем "нижняя цена игры с природой":

$$W = \max_i \min_j W_{ij} \quad (6.17)$$

Правило выбора решения в соответствии с критерием Вальда можно интерпретировать следующим образом: матрица решений $[W_{ir}]$ дополняется еще одним столбцом из наименьших результатов W_{ir} каждой строки. Выбрать надлежит тот вариант, в строке которого стоит наибольшее значение W_{ir} этого столбца [28].

Выбранное таким образом решение полностью исключает риск. Это означает, что принимающий решение не может столкнуться с худшим результатом, чем тот, на который он ориентируется. Какие бы условия V_j не встретились, соответствующий результат не может оказаться ниже W. Это свойство заставляет считать критерий Вальда одним из фундаментальных. Поэтому в технических задачах он применяется чаще всего как сознательно, так и неосознанно. Однако в практических ситуациях излишний пессимизм этого критерия может оказаться очень невыгодным.

Применение этого критерия может быть оправдано, если ситуация, в которой принимается решение, характеризуется следующими обстоятельствами:

- о вероятности появления состояния V_j ничего не известно;
- с появлением состояния V_j необходимо считаться;
- реализуется лишь малое количество решений;
- не допускается никакой риск.

Критерий Байеса-Лапласа в отличие от критерия Вальда, учитывает каждое из возможных следствий всех вариантов решений:

$$W = \max_i \sum_{j=1}^{n} W_{ij} p_i \quad (6.18)$$

Соответствующее правило выбора можно интерпретировать следующим образом: матрица решений $[W_{ij}]$ дополняется еще одним столбцом, содержащим математическое ожидание значений каждой из строк. Выбирается тот вариант, в строках которого стоит наибольшее значение W_{ir} этого столбца.

Критерий Байеса-Лапласа предъявляет к ситуации, в которой принимается решение, следующие требования:

- вероятность появления состояния V_j известна и не зависит от времени;
- принятое решение теоретически допускает бесконечно большое
- количество реализаций;
- допускается некоторый риск при малых числах реализаций.

В соответствии с *критерием Сэвиджа* в качестве оптимальной выбирается такая стратегия, при которой величина риска принимает наименьшее значение в самой неблагополучной ситуации:

$$W = \min_i \max_j (W\max_j - W_{ij}) \quad (6.19)$$

Здесь величину W можно трактовать как максимальный дополнительный выигрыш, который достигается, если в состоянии V_j вместо варианта U_i выбрать другой, оптимальный для этого внешнего состояния, вариант.

Соответствующее критерию Сэвиджа правило выбора следующее: каждый элемент матрицы решений $[W_{ij}]$ вычитается из наибольшего результата max W_{ij} соответствующего столбца. Разности образуют матрицу остатков. Эта матрица пополняется столбцом наибольших разностей W_{ir}. Выбирается тот вариант, в строке которого стоит наименьшее значение.

Согласно *критерию Гурвица* выбирается такая стратегия, которая занимает некоторое промежуточное положение между крайним пессимизмом и оптимизмом:

$$W = \max_j [\alpha \min_i W_{ij} + (1-\alpha) \max_i W_{ij}] \qquad (6.20)$$

где

□ - коэффициент пессимизма, выбираемый в интервале [0,1].

Правило выбора согласно этому критерию следующее: матрица решений $[W_{ij}]$ дополняется столбцом, содержащим средние взвешенные наименьшего и наибольшего результатов для каждой строки (2.6). Выбирается тот вариант, в строках которого стоят наибольшие элементы W_{ir} этого столбца.

При □ =1 критерий Гурвица превращается в критерий Вальда (пессимиста), а при □ =0 - в критерий азартного игрока. Отсюда ясно, какое значение имеет весовой множитель □ . В технических приложениях правильно выбрать этот множитель бывает так же трудно, как правильно выбрать критерий. Поэтому чаще всего весовой множитель □ =0.5 принимается в качестве средней точки зрения.

Критерий Гурвица предъявляет к ситуации, в которой принимается решение, следующие требования:

- о вероятности появления состояния V_j ничего не известно;
- с появлением состояния V_j необходимо считаться;
- реализуется лишь малое количество решений;
- допускается некоторый риск.

Критерий Ходжа-Лемана базируется одновременно на критериях Вальда и Байеса-Лапласа:

$$W = \max_i z \sum_i^n W_{ij} k_j + (1-z) \min_j W_{ij} \quad . (6.20)$$

Правило выбора, соответствующее этому критерию, формулируется следующим образом: матрица решений $[W_{ij}]$ дополняется столбцом, составленным из средних взвешенных (с постоянными весами) математического ожидания и наименьшего результата каждой строки. Отбирается тот вариант решения, в строке которого стоит наибольшее значение этого столбца.

При z=1 критерий преобразуется в критерий Байеса-Лапласа, а при z=0 превращается в критерий Вальда. Таким образом, выбор параметра z подвержен влиянию субъективизма. Кроме того, без внимания остается и число реализаций. Поэтому этот критерий редко применяется при принятии технических решений.

Критерий Ходжа-Лемана предъявляет к ситуации, в которой принимается решение, следующие требования:

- о вероятности появления состояния V_j ничего не известно, но некоторые предположения о распределении вероятностей возможны;
- принятое решение теоретически допускает бесконечно большое количество реализаций; допускается некоторый риск при малых числах реализаций.

Общие рекомендаций по выбору того или иного критерия дать затруднительно. Однако отметим следующее: если в отдельных ситуациях не допустим даже минимальный риск, то следует применять критерий Вальда; если определенный риск вполне приемлем, то можно воспользоваться критерием Сэвиджа. Можно рекомендовать одновременно применять поочередно различные критерии. После этого среди нескольких вариантов, отобранных таким образом в качестве оптимальных, приходится волевым решением выделять некоторое окончательное решение [26,28].

Такой подход позволяет, во-первых, лучше проникнуть во все внутренние связи проблемы принятия решений и, во-вторых, ослабляет влияние субъективного фактора. Кроме того, в области технических задач различные критерии часто приводят к одному результату.

Применение данных критериев с методической точки зрения удобно продемонстрировать на примере одной задачи.

Учет активных условий

Как правило, решение практических задач, связанных с оценкой качества и надежности изделий лесного машиностроения, зависит не только от оперирующей стороны (допустим, конструктора), но и от действий других субъектов системы (например, технолога-лесозаготовителя). Каждая из сторон преследует собственные цели, не всегда совпадающие друг с другом. Неопределенность такого рода при принятии решений относят к классу поведенческих неопределенностей. Теоретической основой нахождения оптимального решения в условиях неопределенности и конфликтных ситуаций является теория игр. Игра - это математическая модель процесса функционирования конфликтующих элементов систем, в котором действия игроков происходят по определенным правилам, называемых *стратегиями*. Ее широкому распространению в последнее время способствовало как развитие ЭВМ, так и создание аналитического аппарата, позволяющего находить аналитические решения для широкого класса задач. Основной постулат теории игр [26,28] - любой субъект системы по меньшей мере так же разумен, как и оперирующая сторона и делает все возможное, чтобы достигнуть своих целей. От реального конфликта игра (математическая модель конфликта) отличается тем, что она ведется по определенным правилам, которые устанавливают порядок и очередность действий субъектов системы, их информированность, порядок обмена информацией, формирование результата игры.

Существует много классов игр, различающихся по количеству игроков, числу ходов, характеру функций выигрыша и т.д. Выделим следующие основные классы игр:

- антагонистические (игры со строгим соперничеством) и неантогонистические. В первом случае цели игроков противоположны, во - втором - могут совпадать;
- стратегические и нестратегические (в первых субъект системы действует независимо от остальных, преследуя свои цели, во-вторых субъекты выбирают единую для всех стратегию);
- парные игры и игры для N-лиц;
- коалиционные и бескоалиционные;
- кооперативные и некооперативные (в первых возможен обмен информацией о возможных стратегиях игроков);
- конечные и бесконечные (в первых - конечное число стратегий).

Наиболее полный обзор направлений теории игр в ее современном состоянии дан в работе [29].

Наибольшее распространение в технических приложениях имеют парные стратегические бескоалиционные конечные некооперативные игры. Модель проблемной ситуации в этом случае имеет вид:

$< U, V, W_1, W_2, R_1, R_2 >$,

где

U - множество стратегий оперирующей стороны (конструктора);

V - множество стратегий оппонирующей стороны (технолог и природа);

W_1 и W_2 - показатели качества игроков;

R_1 и R_2 - системы предпочтения игроков.

Системы предпочтения игроков, в свою очередь, основываются на двух ведущих принципах рационального поведения: принципе наибольшего гарантированного результата и принципе равновесия.

Первый основан на том, что рациональным выбором одного из игроков должен считаться такой, при котором он рассчитывает на самую неблагоприятную для него реакцию со стороны другого игрока.

Второй принцип гласит, что рациональным выбором любого игрока считается такая стратегия $u_\$$ (или $v_\$$), для которой ситуация ($u_\$$, $v_\$$) обоюдовыгодна: любое отклонение от данной ситуации игры не является выгодным ни для одного из игроков.

Решается парная матричная игра (проектируемое изделие - меры и средства противодействия) с нулевой суммой (выигрыш одной стороны равен проигрышу другой) на основе рассмотрения платежной матрицы, которая представляет собой совокупность значений U и V (пара стратегий (u,v) U x V называется *ситуацией игры*) а также выигрышей W_{ij} при парном сочетании всевозможных стратегий сторон.

Решение парной матричной игры может быть в чистых стратегиях, когда для каждой из сторон может быть определена единственная оптимальная стратегия, отклонение от которой невыгодно обоим игрокам. Если выгодно использовать несколько стратегий с определенной частотой их чередования, то решение находится в смешанных стратегиях.

Основные особенности использования методов теории заключаются в следующем. В качестве возможных стратегий со стороны проектируемой системы рассматриваются возможные варианты ее строения, из которых следует выбрать наиболее рациональный. В качестве стратегий противника рассматриваются возможные варианты его противодействия, стратегии их применения.

Необходимо отметить, что при рассмотрении игр с использованием адаптивной системы число ее стратегий может быть существенно расширено благодаря реализации "гибких" конструкторских решений. Анализ игровых ситуаций в этом случае может быть направлен не только на выбор рационального варианта проектируемого изделия, но и на определение алгоритмов рационального применения системы в конфликтной ситуации.

Другая особенность применения методов теории игр заключается в выборе решений, получаемых на основе анализа конфликтной ситуации. В теории игр доказывается теорема о том, что оптимальная стратегия для каждого из игроков является оптимальной и для другого. Так, если решение игры получено в чистых стратегиях (имеется седловая точка), то выбор решения однозначен. Например, если для парной антагонистической игры 3x4 составить матрицу, где элементами u_{ij} будут выигрыши (проигрыши) игроков, то седловая точка находится на пересечении максимина строк и минимакса столбцов

Стратегии	Стратегии В				min
А	1	2	3	4	строк
1	8	2	9	5	2
2	6	5	7	18	***5***
3	7	3	-4	10	-4
max столбцов	8	***5***	9	18	

Оптимальными стратегиями будут для А - 2, для В - 2. Цена игры равна 5. Отметим, что в случае наличия седловой точки ни один из игроков не может улучшить стратегию и стратегии называются *чистыми*. Отметим, что игра с чистыми стратегиями может существовать только при наличии полной информации о действиях противника.

Если же решение игры получено в смешанных стратегиях, то это эквивалентно созданию множества вариантов проектируемого компонента и использованию их с оптимальными частотам, соответствующими оптимальной смешанной стратегии. В случаях, когда не имеется полной информации о действиях противника, вводятся вероятности применения той или иной стратегии в виде векторов

$P_{<n>}=<p_1, p_2, ..., p_n>$ - для игрока А, где $\sum_{i=1}^{n} p_i = 1$;

$Q_{<m>}=<q_1, q_2, ..., q_n>$ - для игрока В, где $\sum_{i=1}^{m} q_i = 1$.

При этом игрок А выбирает стратегию в соответствии с принципом максимина по выражению:

$$\max_{p_i}\{\min(\sum_{i=1}^{n} a_{i1}p_i, \sum_{i=1}^{n} a_{i2} \dots \sum_{i=1}^{n} a_{in}p_i)\},$$

а игра В по принципу минимакса

$$\min_{p_i}\{\max(\sum_{j=1}^{m} a_{j1}q_i, \sum_{j=1}^{m} a_{j2}q_j \dots \sum_{j=1}^{m} a_{jn}q_j)\}.$$

Рассмотрим пример: пусть рассматривается принятие решения в игре 2х2, где игрок А знает вероятность стратегии 1, то есть p_1, тогда очевидно вероятность стратегии 2 будет 1-p, соответственно стратегии игрока В будут q_1 и $1-q_1$. Платежная матрица будет иметь вид:

		B	
		q_1	$1-q_1$
A	p_1	a_{11}	a_{12}
	$1-p_1$	a_{21}	a_{22}

На основании матрицы и приведенных выше выражений составляется таблица:

Чистые стратегии игрока В	Ожидаемые выигрыши игрока А
1	$(a_{11}-a_{21})p_1 + a_{21}$
2	$(a_{12}-a_{22})p_1 + a_{22}$

Из таблицы видно, что ожидаемый выигрыш игрока А линейно зависит от вероятности p_1 (в данном случае задача может быть решена графоаналитически). Тогда смешанная стратегия игрока А будет иметь вид

$<p^*_1, p^*_2>$,

то есть игроку А выгодно применять стратегию 1 с частотой (вероятностью) - p_1, а стратегию 2 с частотой p_2.

Очевидно, что разработка нескольких вариантов изделия сопряжена с большими затратами, не всегда реализуема и затрудняет использование системы. Поэтому при получении решения в смешанных стратегиях рекомендуются следующие случаи принятия окончательного решения [20,26,28]:

- для дальнейшего проектирования выбирается тот вариант, который гарантирует максимальное качество (выбор по максиминной стратегии аналогично критерию Вальда);
- выбирается тот вариант, который в смешанной стратегии должен использоваться с максимальной вероятностью;
- реализуется несколько вариантов изделия с частотами, соответствующими смешанной стратегии (создание адаптивно-модульных конструкций).

Важное значение в задачах исследования качества адаптивных систем имеет не только решение игры, но и анализ платежной матрицы. Это особенно важно в тех случаях, когда решение в смешанных стратегиях не реализуется. Этот анализ может проводиться на основе: оценки возможных потерь эффективности в случае реализации чистой стратегии; определения дополнительных затрат на их компенсацию с помощью "гибких" конструкторских решений; оценки достоверности рассмотренных стратегий противодействия; определения возможности реализации компромиссных вариантов и т.д.

Для анализа конфликтной ситуации требуется на основе математической модели операции построить платежную матрицу $[W_{mn}]$ =$[W_{ij}]$, где W_{ij} характеризует качество изделия при выборе i-го варианта проектируемого изделия и при j-м варианте противодействия противника.

Решение может быть получено в чистых стратегиях, когда есть седловая точка. Условие седловой точки имеет вид

$$\max_i \min_j W_{ij} = \min_j \max_i W_{ij}, \quad (6.21)$$

где левая часть выражения - нижняя цена игры, правая - верхняя цена игры.

Если условие (6.8) не выполняется, то седловая точка отсутствует и требуется реализация смешанной стратегии.

Решение в смешанных стратегиях состоит в реализации чистых стратегий с различными вероятностями, задаваемыми распределением:

для проектируемого изделия в виде вектора-столбца

$G = \{g_i\}$, где $i = 1,2 \ldots m$; $\sum_{i=1}^{m} g_i = 1$;

для противодействия в виде вектора-строки

$F = \{f_j\}$, где $j = 1,2 \ldots n$; $\sum_{j=1}^{n} f_j = 1$,

где

g_i - вероятность выбора стратегии u_i;

f_j - вероятность выбора стратегии v_j.

Платежную функцию запишем в следующем виде:

$$W(G,F) = G^T W_{ij} F^T = \sum_{i=1}^{m}\sum_{j=1}^{n} W_{ij} g_i f_j \qquad (6.22)$$

где индексом "т" обозначена процедура транспонирования.

Платежная функция W(G,F) всегда имеет седловую точку, т.е. всегда существует решение матричной игры. Это утверждение соответствует основной теореме теории матричных игр [26]: каждая матричная игра с нулевой суммой имеет, по крайней мере, одно решение в чистых или смешанных стратегиях.

Последовательность решения игры следующая:

1. Анализируется платежная матрица на предмет исключения заведомо невыгодных и дублирующих стратегий.
2. Проверяется наличие седловой точки по условию (6.21).
3. Если решение в чистых стратегиях отсутствует, то ищется решение в смешанных стратегиях с помощью методов линейного программирования или методом Монте-Карло.

замена	2	6	5	17	18	7	18
(стра-	3	7	3	14	10	8	14
тегии	4	4	6	16	9	19	19
эксплуа-	5	12	4	15	8	10	15
тации)	min столбца	6	2	***9***	5	6	

1.2 Лекция № 2
Тема: «Математические модели операций»

1.2.1 Вопросы лекции:

1. Математическое моделирование в условиях «доброкачественной» и «дурной» неопределенности

1.2.2 Краткое содержание вопросов:

1. Математическое моделирование в условиях «доброкачественной» и «дурной» неопределенности

Реальные задачи исследования операций чаще всего содержат помимо этих двух групп еще одну — неизвестные факторы, которые в совокупности мы обозначим одной буквой . Итак, показатель эффективности W зависит от всех трех групп факторов:

$$W = W(\alpha,\ x,\ \xi).$$

Так как величина W зависит от неизвестных факторов 1, то даже при заданных х она уже не может быть вычислена, остается неопределенной. Задача поиска оптимального решения тоже теряет определенность. Ведь не можем же мы максимизировать неизвестную величину . И все-таки нас не покидает желание сделать эту неизвестную величину по возможности максимальной. Ведь добиваются же успеха люди в условиях, когда не вся обстановка ясна?

Иногда добиваются. Переводя сказанное на математический язык, поставим перед собой следующую задачу.

При заданных условиях а, с учетом неизвестных факторов ?, найти такое решение , которое, по возможности, обеспечивает максимальное значение показателя эффективности W.

Это уже другая, не чисто математическая задача (недаром в ее формулировке сделана оговорка «по возможности»). Наличие неопределенных факторов переводит задачу в новое качество: она превращается в задачу о выборе решения в условиях неопределенности.

Порассуждаем немного о возникшей задаче. Прежде всего, будем честны: неопределенность есть неопределенность, и ничего хорошего в ней нет. Если условия операции неизвестны, мы не можем так же успешно оптимизировать решение, как мы это сделали бы, если бы располагали большей информацией. Поэтому любое решение, принятое в условиях неопределенности, хуже решения, принятого в заранее известных условиях. Что делать? Плохое или хорошее — решение все равно должно быть принято. Наше дело — придать этому решению в возможно большей мере черты разумности. Недаром Т. Л. Саати, один из видных зарубежных специалистов по исследованию операций, определяя свой предмет, говорит не без иронии: «Исследование операций представляет собой искусство давать плохие ответы на практические вопросы, на которые даются еще худшие ответы другими методами».

Задача принятия решения в условиях неопределенности на каждом шагу встречается нам в жизни. Например, мы собрались путешествовать и укладываем в чемодан вещи. Размеры и вес чемодана, а также имеющийся у нас набор вещей заданы (условия а), погода в районе путешествия заранее неизвестна (условия §). Какие предметы одежды Ы надо взять с собой? Эта задача, внешне сходная с задачами исследования операций, конечно, решается нами без всякой математики (незачем «стрелять из пушек по воробьям»), но все же не без опоры на некоторые статистические данные, скажем, о вероятной погоде в районе путешествия, а также собственной склонности к простудам; нечто вроде оптимизации решения, сознательно или бессознательно, мы производим.

Любопытно, что разные люди, при этом, по-видимому, пользуются разными показателями эффективности. Если" молодой человек, скорее всего, стремится максимизировать сумму приятных впечатлений (оставим в стороне вопрос о том, как ее оценить количественно), то пожилой путешественник, пожалуй, предпочтет минимизировать вероятность заболевания...

Теперь возьмем более серьезную задачу. Планируется ассортимент товаров для распродажи на ярмарке. Желательно было бы максимизировать прибыль. Однако заранее неизвестно ни количество покупателей, которые придут на ярмарку, ни потребности каждого из них. Как быть? Неопределенность налицо, а принимать решение нужно!

Другой пример: проектируется система сооружений, оберегающих район от паводков. Ни моменты их наступления, ни размеры заранее неизвестны. А проектировать все-таки нужно, и никакая неопределенность не избавит нас от этой обязанности...

Наконец, еще более сложная задача: разрабатывается план развития вооружения на несколько лет вперед. Неизвестны ни конкретный противник, ни вооружение, которым он будет располагать. А решение принимать надо!

Для того, чтобы такие решения принимать не наобум, по вдохновению, а трезво, с открытыми глазами, современная наука располагает рядом приемов. Каким из них воспользоваться — зависит от того, какова природа неизвестных факторов g, откуда они возникают и кем контролируются. Другими словами, с какого вида неопределенностью мы в данной задаче сталкиваемся?

У читателя может возникнуть вопрос: неужели можно классифицировать неопределенности по «родам» и «сортам»? Оказывается, можно.

Прежде всего, рассмотрим наиболее благоприятный для исследования, так сказать, «доброкачественный» вид неопределенности. Это случай, когда неизвестные факторы § представляют собой обычные объекты изучения теории вероятностей — случайные величины (или случайные функции), статистические характеристики которых нам известны или в принципе могут быть получены к нужному сроку. Такие задачи исследования операций мы будем называть стохастическими задачами, а присущую им неопределенность — стохастической неопределенностью.

Приведем пример стохастической задачи исследования операций. Пусть организуется или реорганизуется работа столовой с целью повысить ее пропускную способность. Нам в точности неизвестно, какое количество посетителей придет в нее за рабочий день, когда именно они будут появляться, какие блюда заказывать и сколько времени будет продолжаться обслуживание каждого из них. Однако характеристики этих случайных величин, если сейчас еще не находятся в нашем распоряжении, могут быть получены статистическим путем.

Другой пример: организуется система профилактического и аварийного ремонта технических устройств с целью уменьшить простои техники за счет неисправностей и ремонтов. Отказы техники, длительности ремонтов и профилактик носят случайный характер. Характеристики всех случайных факторов, входящих в задачу, могут быть получены, если собрать соответствующую статистику.

Рассмотрим более подробно этот «доброкачественный» вид неопределенности. Пусть неизвестные факторы представляют собой случайные величины с какими-то, в принципе известными, вероятностными характеристиками — законами распределения, математическими ожиданиями, дисперсиями и т. п. Тогда показатель эффективности W, зависящий от этих факторов, тоже будет величиной случайной. Максимизировать случайную величину невозможно: при любом решении x она остается случайной, неконтролируемой. Как же быть?

Первое, что приходит в голову: а нельзя ли заменить случайные факторы g их средними значениями (математическими ожиданиями)? Тогда задача становится детерминированной может быть решена обычными методами.

Что и говорить — прием соблазнительный и в некоторых случаях даже оправданный.

Ведь на практике, решая большинство задач физики, механики, техники, мы сплошь рядом им пользуемся, пренебрегая случайностью ряда параметров (теплоемкость, индуктивность, коэффициент трения) и заменяя их средними значениями. Весь вопрос в том, насколько случайны эти параметры: если они мало отклоняются от своих математических ожиданий, так поступать можно и нужно. Так же обстоит дело и в исследовании операций: есть задачи, в которых случайностью можно пренебречь. Например, если мы составляем план снабжения группы предприятий сырьем (см. пример 1 § 1), можно в первом приближении пренебречь, скажем, случайностью фактической производительности источников сырья (если, разумеется, его производство хорошо налажено). Тот же прием — пренебречь случайностью и заменить все входящие в задачу случайные величины их математическими ожиданиями — будет ужо опрометчивым, если влияние случайности на интересующий нас исход операции существенно. Возьмем самый грубый пример: пусть мы ведем обстрел какой-то цели, стремясь во что бы то ни стало попасть в нее. Производится несколько выстрелов. Давайте заменим все случайные координаты точек попадания их математическим ожиданием — центром цели. Получится, что любой выстрел с гарантией попадет в цель, что заведомо неверно. Другой, менее очевидный, пример: планируется работа ремонтной мастерской, обслуживающей автобазу. Пренебрежем случайностью момента появления неисправности (т. е. заменим случайное время безотказной работы машины его математическим ожиданием) и случайностью времени выполнения ремонта. И что же окажется? Мастерская, работа которой спланирована без учета случайности, попросту не будет справляться со своей задачей (в этом мы убедимся в главе 6). Встречаются (и очень часто) операции, в которые случайность входит по существу, и свести задачу к детерминированной не удается.

Итак, рассмотрим такую операцию O, где факторы «существенно случайны» и заметно влияют на показатель эффективности W, который тоже «существенно случаев».

Возникает мысль: надо взять в качестве показателя эффективности среднее значение (математическое ожидание) этой случайной величины и выбрать такое решение x, при котором этот усредненный по условиям показатель обращается в максимум:

$$\overline{W} = M\,[W(\alpha, x, \xi)] \Rightarrow \max.$$

Заметим, что именно так мы поступали в § 2, выбирая в качестве показателя эффективности в задачах, содержащих неопределенность, не просто «доход», а «средний доход», не просто «время», а «среднее время», В большинстве случаев такой подход (мы его назовем «оптимизацией в среднем») вполне оправдан. В самом деле, если мы выберем решение так, чтобы среднее значение показателя эффективности обращалось в максимум, то, безусловно, поступим правильнее, чем если бы выбирали решение наобум.

А как же с элементом неопределенности? Конечно, в какой-то мере он сохраняется. Эффективность каждой отдельной операции, проводимой при конкретных значениях случайных факторов может сильно отличаться от ожидаемой как в большую, так, к сожалению, и в меньшую сторону. Нас может утешить то, что оптимизируя операцию «в среднем», мы в конечном счете после многих ее повторений выиграем больше, чем если бы совсем не пользовались расчетом.

Такая «оптимизация в среднем» очень часто применяется на практике в стохастических задачах исследования операций, и пользуются ею обычно не задумываясь над ее правомочностью. А задуматься надо! Чтобы этот прием был законным, нужно, чтобы операция обладала свойством повторяемости, и «недостача» показателя эффективности в одном случае компенсировалась его «избытком» в другом. Например, если мы предпринимаем длинный ряд однородных операций с целью получить максимальный доход, то доходы от отдельных операций суммируются, «минус» в одном случае покрывается «плюсом» в другом, и все в порядке.

Всегда ли это будет так? Нет, не всегда! Чтобы убедиться в этом, рассмотрим пример. Организуется автоматизированная система управления (АСУ) для службы неотложной медицинской помощи большого города. Вызовы, возникающие в разных районах города в случайные моменты, поступают на центральный пункт управления, откуда они передаются на тот или другой пункт неотложной помощи с приданными ему машинами.

Требуется разработать такое правило (алгоритм) диспетчерской работы АСУ, при котором служба в целом будет функционировать наиболее эффективно. Для этого, прежде всего, надо выбрать показатель эффективности W.

Разумеется, желательно, чтобы время T ожидания врача было минимально. Но это время — величина случайная. Если применить «оптимизацию в среднем», то надо выбрать тот алгоритм, при котором среднее время ожидания минимально. Не так ли?

Оказывается, не совсем так! Беда в том, что времена ожидания врача отдельными больными не суммируются: слишком долгое ожидание одного из них не компенсируется почти мгновенным обслуживанием другого. Выбирая в качестве показателя эффективности среднее время ожидания T, мы рискуем дать предпочтение тому алгоритму, при котором среднее - то время ожидания мало, но отдельные больные могут ожидать врача очень долго! Чтобы избежать таких неприятностей, можно дополнить показатель эффективности добавочным требованием, чтобы фактическое время T ожидания врача было не больше какого-то предельного значения . Поскольку T — величина случайная, нельзя просто потребовать, чтобы выполнялось условие можно только потребовать, чтобы оно выполнялось с очень большой вероятностью, настолько большой, чтобы событие было практически достоверным. Ну что же, назначим какое-то значение , близкое к единице (например, 0,99 или 0,995), настолько близкое, что событие с такой вероятностью можно считать практически достоверным, и потребуем, чтобы условие выполнялось с вероятностью, не меньшей, чем :

$$P(T \leqslant t_0) \geqslant \beta.$$

Введение такого ограничения означает, что из области возможных решений X исключаются решения, ему не удовлетворяющие. Ограничения типа (5.3) называются стохастическими ограничениям и; наличие таких ограничений сильно усложняет задачу оптимизации.

Особенно осторожным надо быть с «оптимизацией в среднем», когда речь идет не о повторяемой, массовой операции, а о единичной, «уникальной».

Все зависит от того, к каким последствиям может привести неудача данной операции, т. е. слишком малое значение показателя эффективности W; иногда оно может означать попросту катастрофу. Что толку в том, что операция в среднем приносит большой выигрыш, если в данном, единичном случае она может нас дотла разорить? От таких катастрофических результатов можно опять-таки спасаться введением стохастических ограничений. При достаточно большом значении уровня доверия можно быть практически уверенным в том, что угрожающее разорение нас не постигнет.

Итак, мы вкратце рассмотрели случай «доброкачественной» (стохастической) неопределенности и в общих чертах осветили вопрос об оптимизации решения в таких задачах. Но это, как говорится, цветочки, ягодки будут впереди! Стохастическая неопределенность — это почти определенность, если только известны вероятностные характеристики входящих в задачу случайных факторов. Гораздо хуже обстоит дело, когда неизвестные факторы не могут быть изучены и описаны статистическими методами. Это бывает в двух случаях: либо а) распределение вероятностей для параметров в принципе существует, но к моменту принятия решения не может быть получено, либо б) распределение вероятностей для параметров вообще не существует.

Пример ситуации типа а): проектируется информационно-вычислительная система (ИВС), предназначенная для обслуживания каких-то случайных потоков требований (запросов). Вероятностные характеристики этих потоков требований в принципе могли бы быть получены из статистики, если бы данная ИВС (или аналогичная ей) уже существовала и функционировала достаточно долгое время. Но к моменту создания проекта такой информации нет, а решение принимать надо! Как быть?

Разумеется, можно заранее (из умозрительных соображений) задаться какими-то характеристиками случайных факторов , оптимизировать решение х на этой основе (просто «в среднем» или при стохастических ограничениях) и остановиться на нем.

Это будет, безусловно, лучше, чем выбрать решение наобум, но не намного лучше. Гораздо разумнее будет применить следующий прием: оставить некоторые элементы решения х свободными, изменяемыми. Затем выбрать для начала какой-то вариант решения, зная заведомо, что он не самый лучший, и пустить систему в ход, а потом, по мере накопления опыта, целенаправленно изменять свободные параметры решения, добиваясь того, чтобы эффективность не уменьшалась, а увеличивалась. Такие совершенствующиеся в процессе применения алгоритмы управления называются адаптивными. Преимущество адаптивных алгоритмов в том, что они не только избавляют нас от предварительного сбора статистики, по и перестраиваются в ответ на изменение обстановки. По мере накопления опыта такой алгоритм постепенно улучшается, подобно тому, как живой человек «учится на ошибках».

Теперь обратимся к самому трудному и неприятному случаю б), когда у неопределенных факторов вообще не существует вероятностных характеристик; другими словами, когда их нельзя считать «случайными» в обычном смысле слова.

— Как? — может быть, спросит читатель. — Разве не всякая неопределенность есть случайность?

— Нет, — ответим мы, — лучше пепутать эти понятия.

Поясняем: под термином «случайное явление» в теории вероятностей принято понимать явление, относящееся к классу повторяемых и, главное, обладающее свойством статистической устойчивости. При повторении однородных опытов, исход которых случаен, их средние характеристики проявляют тенденцию к устойчивости, стабилизируются. Частоты событий приближаются к их вероятностям, средние арифметические — к математическим ожиданиям. Если много раз бросать монету, частота появления герба постепенно стабилизируется, перестает быть случайной; если много раз взвешивать на аналитических весах одно и то же тело, средний результат перестает колебаться, выравнивается. Это пример доброкачественной, стохастической неопределенности.

Однако бывает неопределенность и нестохастического вида, которую мы условно назовем «дурной неопределенностью». Несмотря на то, что факторы заранее неизвестны, не имеет смысла говорить об их «законах распределения» или других вероятностных характеристиках.

Пример: допустим, планируется некая торгово-промышленная операция, успех которой зависит от того, юбки какой длины § будут носить женщины через два года. Распределение вероятностей для величины в принципе не может быть получено ни из каких статистических данных. Даже если рассмотреть великое множество опытов (годов), начиная с тех отдаленных времен, когда женщины впервые падели юбки, и в каждом из них зарегистрировать величину это вряд ли поможет нам в нашем прогнозе. Вероятностное распределение величины попросту не существует, так как не существует массива однородных опытов, где она обладала бы должной устойчивостью. Налицо случай «дурной неопределенности».

Как же все-таки быть в таких случаях? Отказаться вообще от применения математических методов и выбрать решение «волевым» образом? Нет, этого делать не стоит. Некоторую пользу предварительные расчеты могут принести даже в таких скверных условиях.

Давайте поразмышляем на эту тему. Пусть выбирается решение в какой-то операции О, условия которой содержат «дурную неопределенность» — параметры относительно которых никаких сведений мы не имеем, а можем делать лишь предположения. Попробуем все же решить задачу.

Первое, что приходит в голову: а ну-ка, зададимся какими-то, более или менее правдоподобными, значениями параметров Тогда задача перейдет в категорию детерминированных и может быть решена обычными методами (детерминированный случай, со всеми его трудностями, представляется нам теперь чуть ли не «курортом».

Нет, радоваться рано. Допустим, что, затратив много усилий и времени (своего и машинного), мы это сделали. Ну и что? Будет ли найденное решение хорошим для других условий ? Как правило, нет. Поэтому ценность его — сугубо ограниченная. В данном случае разумно будет выбрать не решение оптимальное для каких-то условий а некое компромиссное решение, которое, не будучи оптимальным, может быть, ни для каких условий, будет все же приемлемым в целом их диапазоне.

В настоящее время полноценной научной теория компромисса не существует, хотя некоторые попытки в этом направлении в теории игр и статистических решений делаются (см. главу 8). Обычно окончательный выбор компромиссного решения осуществляется человеком. Опираясь на предварительные расчеты, в ходе которых решается большое число прямых задач исследования операций для разных условий и разных вариантов решения х, он может оценить сильные и слабые стороны каждого варианта и на этой основе сделать выбор. Для этого необязательно (хотя иногда и любопытно) здать точный «условный» оптимум для каждой совокупности условий g. Математические вариационные методы в данном случае отступают на задний план.

Подчеркнем еще одну полезную функцию предварительных математических расчетов в задачах с «дурной неопределенностью»: они помогают заранее отбросить те решения , которые при любых условиях уступают другим, т. е. оказываются неконкурентоспособными. В ряде случаев это помогает существенно сузить множество X, иногда — свести его к небольшому числу вариантов, которые легко могут быть просмотрены и оценены человеком в поисках удачного компромисса.

При рассмотрении задач исследования операций с «дурной неопределенностью» всегда полезно сталкивать в споре разные подходы, разные точки зрения. Среди последних надо отметить одну, часто применяемую в силу своей математической определенности, которую можно назвать «позицией крайнего пессимизма». Она сводится к тому, что, принимая решение в условиях «дурной неопределенности», надо всегда рассчитывать на худшее и принимать то решение, которое дает максимальный эффект в наихудших условиях. Если в этих условиях мы получаем выигрыш , то можно гарантировать, что в любых других он будет не меньше («принцип гарантированного результата»). Этот подход привлекателен тем, что дает четкую постановку задачи оптимизации и возможность ее решения корректными математическими методами.

Но он оправдан далеко не всегда. Область его применения — по преимуществу так называемые «конфликтные ситуации», в которых условия зависят от сознательно действующего лица («разумного противника»), отвечающего на любое наше решение наихудшим для нас образом. В более нейтральных ситуациях принцип «гарантированого выигрыша» не является единственно возможным, но может быть рассмотрен наряду с другими. Пользуясь им, нельзя забывать, что эта точка зрения — крайняя, что на ее основе можно выбрать только очень осторожное, «перестраховочное» решение, которое не всегда будет разумным. Вообразите себе, например, военачальника, который всякое свое решение будет принимать исходя из гипотезы, что его противник необычайно умен, хитер и изворотлив и на каждое его действие немедленно ответит наихудшим для него образом. Вряд ли такому военачальнику будет сопутствовать удача! Напротив, в любой конкретной ситуации нужно стараться угадать, в чем слаб и «глуп» противник, и стараться «обвести его вокруг пальца». Тем менее уместен крайне пессимистический подход в ситуациях, где стороне, принимающей решение, не противостоят никакие враждебные силы. Расчеты, основанные на точке зрения «крайнего пессимизма», всегда должны корректироваться разумной долей оптимизма. Вряд ли стоит становиться и на противоположную точку зрения — крайнего или «залихватского» оптимизма, но известная доля риска при принятии решения все же должна присутствовать. Нельзя также забывать о том, что любое решение, принятое в условиях «дурной неопределенности», — неизбежно плохое решение, и вряд ли стоит обосновывать его с помощью тонких и кропотливых расчетов. Скорее следует подумать о том, откуда можно было бы взять недостающую информацию. Здесь все способы хороши — лишь бы прояснить положение.

Упомянем в этой связи об одном довольно оригинальном методе, не пользующемся любовью в среде чистых математиков, но тем не менее полезном, а иногда — единственно возможном.

Речь идет о так называемом методе экспертных оценок. Он часто применяется в задачах, связанных с прогнозированием в условиях дурной неопределенности (например, в футурологии). Грубо говоря, идея метода сводится к следующему: собирается коллектив сведущих, компетентных в данной области людей, и каждому из них предлагается ответить на какой-то вопрос (например, назвать срок, когда будет совершено то или другое открытие, или оценить вероятность того или другого события). Затем полученные ответы обрабатываются наподобие статистического материала. Результаты обработки, разумеется, сохраняют субъективный характер, но в гораздо меньшей степени, чем если бы мнение высказывал один эксперт («ум хорошо, а два — лучше»). Подобного рода экспертные оценки для неизвестных условий могут быть применены и при решении задач исследования операций с дурной неопределенностью. Каждый из экспертов на глаз оценивает степень правдоподобия различных вариантов условий , приписывая им какие-то субъективные вероятности. Несмотря на субъективный характер оценок каждого эксперта, усредняя оценки целого коллектива, можно получить нечто более объективное и полезное (кстати, оценки разных экспертов расходятся не так сильно, как можно было бы ожидать). Таким образом, задача с дурной неопределенностью как бы сводится к обычной стохастической задаче. Разумеется, к полученным отсюда выводам не надо относиться слишком доверчиво, но наряду с другими результатами, вытекающими из других точек зрения, они все же могут помочь при выборе решения.

Наконец, сделаем одно общее замечание. При обосновании решения в условиях неопределенности, что бы мы ни делали, элемент неопределенности, «гадательности» сохраняется. Поэтому нельзя предъявлять к точности решений слишком высокие требования. Вместо того, чтобы указать одно-единственное, в точности «оптимальное» (с какой-то точки зрения) решение, лучше выделить целую область «приемлемых» решений, которые оказываются несущественно хуже других, какой бы точкой зрения мы ни пользовались.

В пределах этой области и должны производить свой окончательный выбор ответственные за это люди. Исследователь, предлагая им рекомендации по выбору решения, всегда должен одновременно указывать точки зрения, из которых вытекают те или другие рекомендации.

1.3 Лекция № 3
Тема: «Задачи теории массового обслуживания»

1.3.1 Вопросы лекции:

1. Основные определения в территории графов и в сетевых задачах (граф ориентированный, граф и ориентированный цикл, дерево, цепь, путь, контур, петля).

1.3.2 Краткое содержание вопросов:

1. Основные определения в территории графов и в сетевых задачах (граф ориентированный, граф и ориентированный цикл, дерево, цепь, путь, контур, петля).

Если рассматривать логистику как задачу нахождения оптимального перемещения, то с этой точки зрения – каждый человек в течении дня на уровне подсознания неоднократно решает такую задачу, перемещаясь:
– из дома на работу и обратно;
– с работы в театр, в кино, в кафе и т.п.;
– в столовую во время обеда, по помещениям и учреждениям во время работы, и т.д.
– даже первобытный человек во время охоты, решал данную задачу в процессе погони за добычей!
В своей общественно-практической деятельности человек также каждодневно решает задачи логистики оптимального перемещения:
– грузов;
– войск;
– готовой продукции и комплектующих в процессе её производства;
– пассажиров и многого другого.

Логистика – как наука известна со времен Римской империи, где: “Логист” – снабженец в армии.

В настоящее время – основная задача логистики, это:

Линейная транспортная задача нахождения способов и путей наиболее оптимальной и быстрой доставки грузов, товаров, пассажиров и т.п. к пунктам назначения.*

К началу 20-го века Логистика, как наука, получила в арсенал своих научных средств новый математический аппарат – теорию графов

* Данное определение логистики не охватывает все её задачи и выбрано как наиболее простое, поскольку к настоящему времени дано более 200 определений, многие из которых противоречивы к другим.

2. История возникновения теории графов

Датой рождения теории графов принято считать 1736 год, когда Леонард Эйлер сформулировал первые теоремы новой теории, решая задачу "О кенигсбергских мостах", и написал об этом в письме итальянскому математику и инженеру Марионни 13 марта 1736 г., чем и закрепил за собой приоритет автора новой теории.
Однако: Термин «граф» впервые ввел в математику венгерский математик Денеш Кениг в 1936 году.

В настоящее время:

– в некоторых отраслях прикладной деятельности вместо понятия "граф" используется термин "сеть": сетевой график; сеть железных дорог и т.п.;
– вместо термина "вершина графа" при использовании "сетевой" терминологии используется термин "сетевой узел" или просто – "узел".

"Задача о кенигсбергских мостах".

В 1736 г. город Калининград назывался Кёнигсбергом. Он был расположен на берегах р. Прегель, и на двух островах между ними. Четыре образовавшихся участка суши (правый и левый берег, и два острова) соединяло семь мостов так (рис.1). В задаче предлагалось составить маршрут для почтальона, полицейского или туриста, чтобы они побывали во всех районах города, и при этом прошли по каждому мосту только один раз – в настоящее время такой путь называется "Эйлеровым путём".

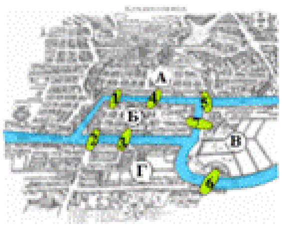
В числе сформулированных Эйлером теорем была теорема о том, что: **Граф с более чем двумя нечётными степенями вершин невозможно начертить или обойти все его вершины так, чтобы пройти по каждому ребру графа только один раз.**

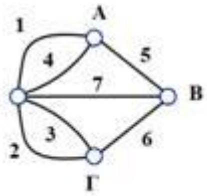
При этом степень вершины он определил равной числу линий – "ребер", связанных с данной вершиной или "сетевым узлом" графа.

На рис. 2 показано изображение графа, для "Задачи о кенигсбергских мостах".

Вершины графа соответствуют определённому району города, а ребра – мостам через реку.
Из рисунка видно, что все четыре вершины графа – имеют нечетные степени: вершины А, В, Г – имеют степень 3, вершина Б – степень 5, т.е. "Эйлерового пути" у данного графа не существует!

На рис. 3 приведены примеры изображений графов, для которых существуют "Эйлеровы пути".





Эйлеровы Графы:

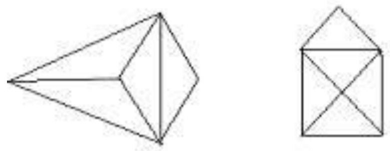


Рис. 1. Рис. 2.

Рис. 3.

Далее, вплоть до начала 20-го века теория графов развивалась в основном в виде формулирования новых теорем, сформированных по результатам решения различных "головоломных задач".

Серьезное развитие теория графов получила в связи с возникновением массового крупносерийного производства, общим всплеском науки и технологий в первой половине 20-го века.

3. Области применения графов.

В настоящее время теория графов нашла очень широкое применение в:

1. В микроэлектронике – при разработке топологии микросхем.
2. При разработке сложных электрических схем и схем их монтажа в электротехнических шкафах, в электрощитах.
3. В химии – при разработке новых сложных молекулярных соединений и т.п.
4. Физике – при описании и анализе схем развития квантовых процессов.
5. При разработке коммуникационных систем различного назначения.
6. В транспортных системах – как для изучения самих систем, так и при составлении оптимальных маршрутов доставки грузов – логистика.
7. В информатике и программирование – при разработке алгоритмов расчетов, программ.
8. В экономике и планировании – в виде сетевых графиков.
и т.д.

4. Основные понятия теории графов.

Графом, или точнее: "**Изображением графа**" – называется конечное множество точек, соединенных, **как правило**, друг с другом любыми линиями. Данные точки называются вершинами графа или узлами сети, а соединяющие их линии – рёбрами. Если, точка не соединена ни с одной другой точкой – она называется "**висячей**". **Степень вершины** графа равна количеству линий или ребер, связанных с данной точкой, вершиной, узлом. p(N)
С точки зрения решения задачи, связанной с конкретным изображением графа: **Граф** непосредственнл **– это решение поставленной задачи** в виде конкретной схемы пути по вершинам и ребрам изображения графа. Каждая задача может иметь одно и более решений, либо не иметь решения.
Эйлеров граф – это такой граф, в котором существует путь по всем его ребрам такой, чтобы каждое ребро было пройдено только один раз.

Гамильтонов путь – путь через все вершины графа наиболее коротким путем. Пути могут быть замкнутыми, т.е. начинаться и заканчиваться в одной точке, либо разомкнутыми.
Ориентированные графы – граф, имеющий ребра по которым можно перемещаться только в одном направлении.
Неориентированные графы – в котором по всем ребрам можно перемещаться в любую сторону.

Элементарные теоремы:

$$\sum_{i=1}^{Все} p(i) = 2R$$

Теорема 1.

Сумма степеней вершин графа – четное и равна удвоенному числу ребер графа.
Доказательство: Каждое ребро соединяет две вершины (например А и В), и будет дважды учтено при определении степеней узлов, поскольку: p(A) = p(B) =1 и p(A) + p(B) = 2.

Теорема 2.

Число узлов с нечетной степенью любого графа – четное.

Доказательство: Из Теоремы 1 мы знаем, что сумма четных и нечетных степеней вершин – четная: $\sum p_ч + \sum p_н = 2R$, а сумма четных чисел $\sum p_ч$ – всегда четная. Значит и $\sum p_н$ – четная.

Теорема 3.

У Графа с количеством вершин нечетной степени больше 2-х Эйлерова пути нет.

Доказательство: выполнено Эйлером в 1736 г.

5. Наиболее известные и "яркие" задачи теории графов:

1. **Теорема о четырёх красках** утверждает, что всякую расположенную на сфере карту можно раскрасить четырьмя красками так, чтобы любые две области, имеющие общий участок границы, были раскрашены в разные цвета. Теорема была сформулирована в 1852 году, однако доказать ее долгое время не удавалось. В течение этого времени было предпринято множество попыток как доказательства, так и опровержения, и эта задача носила название *проблемы четырёх красок*. Для простых карт достаточно и трёх цветов, а четвёртый цвет начинает требоваться, например, тогда, когда имеется одна область, окруженная нечетным числом других, которые соприкасаются друг с другом, образуя цикл. **Теорема о пяти красках**, утверждающая, что достаточно пяти цветов, имела короткое несложное доказательство и была доказана в конце XIX века, но доказательство теоремы для случая четырёх цветов столкнулось со значительными трудностями и была доказана лишь в 1976 году. Это была первая крупная математическая теорема, доказанная с помощью компьютера. Доказательство этого факта заняло сотни страниц и не все

математики признали его. В 1997 году, и в 2005 году были найдены более простые доказательства, также проделанные с помощью компьютера.

2. **Задача коммивояжера** – в которой необходимо посетить каждый город в пределах определенной территории и возвратиться в пункт отправления, причем так, чтобы путь был как можно короче. С современной точки зрения – это типичная задача логистики, была придумана исключительно для развлечения.

3. **Задача о ходе коня** – задача о нахождении маршрута шахматного коня, проходящего через все поля стандартной шахматной доски по одному разу. Как оказалось: Количество всех замкнутых маршрутов коня (гамильтоновых циклов) без учёта направления обхода равно:
13 267 364 410 532.Количество всех незамкнутых маршрутов (с учётом направления обхода) равно: 19 591 828 170 979 904. Наиболее красивое решение (рис. 4) получено при помощи "Шахматного компьютера". Однако с точки зрения логистики наиболее интересным кажется маршрут, найденный К. Я. Янишем, в котором конь сначала обходит одну половину доски, а затем вторую.

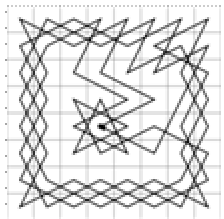


Рис. 4.

6. Применение Теории Графов в логистике.

Как уже указано выше – нахождение замкнутого Гамильтонова пути – это фактически и есть **главная задача современной логистики**.

1.4 Лекция № 4
Тема: «Статистическое моделирование случайных процессов»

1.4.1 Вопросы лекции:

1. Методы моделирования случайных процессов.

2. Применимость различных методов в зависимости от вида условий – ограничений.

1.4.2. Краткое содержание вопросов:

1. Методы моделирования случайных процессов

Большой класс СП, имеющих место в информационно-измерительных системах, системах автоматического управления, а также в каналах связи, подверженных воздействию случайных возмущений описывается с помощью ДУ вида

$$\dot{\mathbf{y}} = f\left(\mathbf{y}, t, \mathbf{x}(t)\right),\ \mathbf{y}(0) = \mathbf{y}_0,\ t \in [0, T], \qquad (3.25)$$

где $\mathbf{y} = (y_1, \ldots, y_n)^T$ - вектор состояния системы, $\mathbf{x}(t) = \left(x_1(t), \ldots, x_m(t)\right)^T$ - векторный стационарный СП.

При исследовании на ЭВМ системы (3.25) необходимо получать реализации СП. Методы моделирования СВ рассматривались в главе 2. Ниже приводятся некоторые распространенные на практике методы моделирования гауссовских стационарных СП: метод формирующего фильтра (п. 3.5.1), метод скользящего суммирования (п. 3.5.2) и рекуррентные моделирующие алгоритмы (п. 3.5.3).

При имитации системы (3.25) на ЭВМ осуществляется переход от непрерывной системы к ее дискретной модели. Как правило, используются численные методы, входящие в математическое обеспечение ЭВМ (например, метод Рунге-Кутта и его модификации) [37]. При этом возникают методические ошибки, в том числе и при получении реализаций СП. Величина ошибок определяется выбранным шагом интегрирования h.

Для линейных стационарных систем, находящихся под воздействием гауссовских стационарных случайных возмущений, могут быть получены алгоритмы моделирования, лишенные методических ошибок. Эти алгоритмы рассматриваются в п. 3.5.3. Они основаны на методе дискретизации линейных стохастических уравнений . Метод дискретизации дает сравнительно простые и легко реализуемые алгоритмы моделирования гауссовских векторных и скалярных СП с дробно-рациональным спектром высокого порядка. В п. 3.5.4 метод дискретизации применяется для процессов с типовыми КФ. Помимо задач цифрового моделирования алгоритмы дискретизации оказываются полезными при расчетах корреляционных характеристик линейных систем и применении методов оптимальной фильтрации к обработке СП.

В настоящее время разработан ряд методов моделирования гауссовских стационарных СП $\mathbf{x}(t)$ с заданными характеристиками: математическим ожиданием m_x, КФ $R_x(\tau)$ или спектральной плотностью $S_x(\omega)$. При решении задач моделирования в целях удобства зачастую считают математическое ожидание нулевым, а дисперсию σ_x^2 - единичной.

Использование преобразования

$$x(t) = m_x + \sigma_x \overset{\circ}{x}(t)$$

позволяет получить процессы с требуемыми значениями этих характеристик. Здесь $M\left[\overset{\circ}{x}(t)\right] = 0$, $\sigma^2\left[\overset{\circ}{x}(t)\right] = 1$. Как известно, любая из функций $R_x(\tau)$ или $S_x(\omega)$ описывает полностью рассматриваемый класс процессов. Обе характеристики связаны взаимно однозначно преобразованиями Фурье [36]:

$$R_x(\tau) = \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau}\, d\omega, \qquad (3.26)$$

$$S_x(\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}R_x(\tau)\,e^{-j\omega\tau}\,d\tau \qquad (3.27)$$

Обычно задача формулируется следующим образом. По известным характеристикам процесса (математическому ожиданию, дисперсии и КФ или спектральной плотности) требуется построить вычислительный алгоритм, позволяющий получать на ЭВМ реализации СП $x(t)$ или последовательностей $x_k,\ k=0,1,\ 2,\dots$. В гауссовском случае модель процесса, заданная математическим ожиданием и КФ, является полностью определенной. Шаг дискретизации Δt может быть не равен шагу интегрирования h системы (3.25).

Известные методы можно разбить на две большие группы: **точные** (метод рекуррентных алгоритмов дискретизации) и **приближенные** (методы формирующего фильтра, скользящего суммирования). В точных методах отсутствует методическая ошибка по КФ, т. е. КФ $R_x[l]=M[x_{k+l}x_k]$ последовательности x_k равна дискретным значениям $R_x(k\Delta t)$ КФ $R_x(\tau)$ моделируемого процесса с непрерывным временем [41].

Для приближенных методов равенство заданных и воспроизводимых на ЭВМ характеристик выдерживается не точно, с некоторой погрешностью [41]. В настоящее время практически отсутствуют работы по анализу погрешностей приближенных методов моделирования, поэтому основным и наиболее надежным способом контроля приближенных алгоритмов является статистическая обработка моделируемых реализаций.

2. Применимость различных методов в зависимости от вида условий – ограничений.

Объектом исследования в прикладной статистике являются статистические данные, полученные в результате наблюдений или экспериментов. Статистические данные – это совокупность объектов (наблюдений, случаев) и признаков (переменных), их характеризующих. Например, объекты исследования – страны мира и признаки, – географические и экономические показатели их характеризующие: континент; высота местности над уровнем моря; среднегодовая температура; место страны в списке по качеству жизни, доли ВВП на душу населения; расходы общества на здравоохранение, образование, армию; средняя продолжительность жизни; доля безработицы, безграмотных; индекс качества жизни и т.д.
Переменные – это величины, которые в результате измерения могут принимать различные значения.
Независимые переменные – это переменные, значения которых в процессе экперимента можно изменять, а зависимые переменные – это переменные, значения которых можно только измерять.
Переменные могут быть измерены в различных шкалах. Различие шкал определяется их информативностью. Рассматривают следующие типы шкал, представленные в порядке возрастания их информативности: номинальная, порядковая, интервальная, шкала отношений, абсолютная. Эти шкалы отличаются друг от друга также и количеством допустимых математических действий. Самая «бедная» шкала – номинальная, так как не

определена ни одна арифметическая операция, самя «богатая» – абсолютная. Измерение в номинальной (классификационной) шкале означает определение принадлежности объекта (наблюдения) к тому или иному классу. Например: пол, род войск, профессия, континент и т.д. В этой шкале можно лишь посчитать количество объектов в классах – частоту и относительную частоту.
Измерение в порядковой (ранговой) шкале, помимо определения класса принадлежности, позволяет упорядочить наблюдения, сравнив их между собой в каком-то отношении. Однако эта шкала не определяет дистанцию между классами, а только то, какое из двух наблюдений предпочтительнее. Поэтому порядковые экспериментальные данные, даже если они изображены цифрами, нельзя рассматривать как числа и выполнять над ними арифметические операции 5 . В этой шкале дополнительно к подсчету частоты объекта можно вычислить ранг объекта. Примеры переменных, измеренных в порядковой шкале: бальные оценки учащихся, призовые места на соревнованиях, воинские звания, место страны в списке по качеству жизни и т.д. Иногда номинальные и порядковые переменные называют категориальными, или группирующими, так как они позволяют произвести разделение объектов исследования на подгруппы.
При измерении в интервальной шкале упорядочивание наблюдений можно выполнить настолько точно, что известны расстояния между любыми двумя их них. Шкала интервалов единственна с точностью до линейных преобразований ($y = ax + b$). Это означает, что шкала имеет произвольную точку отсчета – условный нуль. Примеры переменных, измеренных в интервальной шкале: температура, время, высота местности над уровнем моря. Над переменными в данной шкале можно выполнять операцию определения расстояния между наблюдениями. Расстояния являются полноправными числами и над ними можно выполнять любые арифметические операции.
Шкала отношений похожа на интервальную шкалу, но она единственна с точностью до преобразования вида $y = ax$. Это означает, что шкала имеет фиксированную точку отсчета – абсолютный нуль, но произвольный масштаб измерения. Примеры переменных, измеренных в шкале отношений: длина, вес, сила тока, количество денег, расходы общества на здравоохранение, образование, армию, средняя продолжительность жизни и т.д. Измерения в этой шкале – полноправные числа и над ними можно выполнять любые арифметические действия.
Абсолютная шкала имеет и абсолютный нуль, и абсолютную единицу измерения (масштаб). Примером абсолютной шкалы является числовая прямая. Эта шкала безразмерна, поэтому измерения в ней могут быть использованы в качестве показателя степени или основания логарифма. Примеры измерений в абсолютной шкале: доля безработицы; доля безграмотных, индекс качества жизни и т.д.
Большинство статистических методов относятся к методам параметрической статистики, в основе которых лежит предположение, что случайный вектор переменных образует некоторое многомерное распределение, как правило, нормальное или преобразуется к нормальному распределению. Если это предположение не находит подтверждения, следует воспользоваться непараметрическими методами математической статистики.

Корреляционный анализ. Между переменными (случайными величинами) может существовать функциональная связь, проявляющаяся в том, что одна из них определяется как функция от другой. Но между переменными может существовать и связь другого рода, проявляющаяся в том, что одна из них реагирует на изменение другой изменением своего закона распределения. Такую связь называют стохастической. Она появляется в том случае, когда имеются общие случайные факторы, влияющие на обе переменные. В качестве меры зависимости между переменными используется коэффициент корреляции (r), который изменяется в пределах от –1 до +1. Если коэффициент корреляции отрицательный, это означает, что с увеличением значений одной переменной значения другой убывают. Если переменные независимы, то коэффициент корреляции равен 0 (обратное утверждение верно только для переменных, имеющих нормальное распределение). Но если коэффициент корреляции не равен 0 (переменные называются некоррелированными), то это значит, что между переменными существует зависимость. Чем ближе значение r к 1, тем зависимость сильнее. Коэффициент корреляции достигает своих предельных значений +1 или -1, тогда и только тогда, когда зависимость между переменными линейная. Корреляционный анализ позволяет установить силу и направление стохастической взаимосвязи между переменными (случайными величинами). Если переменные измерены, как минимум, в интервальной шкале и имеют нормальное распределение, то корреляционный анализ осуществляется посредством вычисления коэффициента корреляции Пирсона, в противном случае используются корреляции Спирмена, тау Кендала, или Гамма.

Регрессионный анализ. В регрессионном анализе моделируется взаимосвязь одной случайной переменной от одной или нескольких других случайных переменных. При этом, первая переменная называется зависимой, а остальные – независимыми. Выбор или назначение зависимой и независимых переменных является произвольным (условным) и осуществляется исследователем в зависимости от решаемой им задачи. Независимые переменные называются факторами, регрессорами или предикторами, а зависимая переменная – результативным признаком, или откликом.
Если число предикторов равно 1, регрессию называют простой, или однофакторной, если число предикторов больше 1 – множественной или многофакторной. В общем случае регрессионную модель можно записать следующим образом:

$$y = f(x_1, x_2, \dots, x_n),$$

где y – зависимая переменная (отклик), x_i (i = 1,…, n) – предикторы (факторы), n – число предикторов.

Посредством регрессионного анализа можно решать ряд важных для исследуемой проблемы задач:

1). Уменьшение размерности пространства анализируемых переменных (факторного пространства), за счет замены части факторов одной переменной – откликом. Более полно такая задача решается факторным анализом.

2). Количественное измерение эффекта каждого фактора, т.е. множественная регрессия, позволяет исследователю задать вопрос (и, вероятно, получить ответ) о том, «что является лучшим предиктором для...». При этом, становится более ясным воздействие отдельных факторов на отклик, и исследователь лучше понимает природу изучаемого явления.

3). Вычисление прогнозных значений отклика при определенных значениях факторов, т.е. регрессионный анализ, создает базу для вычислительного эксперимента с целью

получения ответов на вопросы типа «Что будет, если… ».
4). В регрессионном анализе в более явной форме выступает причинно-следственный механизм. Прогноз при этом лучше поддается содержательной интерпретации.

Канонический анализ. Канонический анализ предназначен для анализа зависимостей между двумя списками признаков (независимых переменных), характеризующих объекты. Например, можно изучить зависимость между различными неблагоприятными факторами и появлением определенной группы симптомов заболевания, или взаимосвязь между двумя группами клинико-лабораторных показателей (синдромов) больного. Канонический анализ является обобщением множественной корреляции как меры связи между одной переменной и множеством других переменных. Как известно, множественная корреляция есть максимальная корреляция между одной переменной и линейной функцией других переменных. Эта концепция была обобщена на случай связи между множествами переменных – признаков, характеризующих объекты. При этом достаточно ограничиться рассмотрением небольшого числа наиболее коррелированных линейных комбинаций из каждого множества. Пусть, например, первое множество переменных состоит из признаков $y1, \ldots, yp$, второе множество состоит из – $x1, \ldots, xq$, тогда взаимосвязь между данными множествами можно оценить как корреляцию между линейными комбинациями $a1y1 + a2y2 + ... + apyp$, $b1x1 + b2x2 + ... + bqxq$,, которая называется канонической корреляцией. Задача канонического анализа в нахождении весовых коэффициентов таким образом, чтобы каноническая корреляция была максимальной.

Методы сравнения средних. В прикладных исследованиях часто встречаются случаи, когда средний результат некоторого признака одной серии экспериментов отличается от среднего результата другой серии. Так как средние это результаты измерений, то, как правило, они всегда различаются, вопрос в том, можно ли объяснить обнаруженное расхождение средних неизбежными случайными ошибками эксперимента или оно вызвано определенными причинами. Если идет речь о сравнении двух средних, то можно применять критерий Стьюдента (t-критерий). Это параметрический критерий, так как предполагается, что признак имеет нормальное распределение в каждой серии экспериментов. В настоящее время модным стало применение непараметрических критериев сравнения средних
Сравнение средних результата один из способов выявления зависимостей между переменными признаками, характеризующими исследуемую совокупность объектов (наблюдений). Если при разбиении объектов исследования на подгруппы при помощи категориальной независимой переменной (предиктора) верна гипотеза о неравенстве средних некоторой зависимой переменной в подгруппах, то это означает, что существует стохастическая взаимосвязь между этой зависимой переменной и категориальным предиктором. Так, например, если установлено, что неверна гипотеза о равенстве средних показателей физического и интеллектуального развития детей в группах матерей, куривших и не куривших в период беременности, то это означает, что существует зависимость между курением матери ребенка в период беременности и его интеллектуальным и физическим развитием.
Наиболее общий метод сравнения средних дисперсионный анализ. В терминологии дисперсионного анализа категориальный предиктор называется фактором.

Дисперсионный анализ можно определить как параметрический, статистический метод, предназначенный для оценки влияния различных факторов на результат эксперимента, а также для последующего планирования экспериментов. Поэтому в дисперсионном анализе можно исследовать зависимость количественного признака от одного или нескольких качественных признаков факторов. Если рассматривается один фактор, то применяют однофакторный дисперсионный анализ, в противном случае используют многофакторный дисперсионный анализ.

Частотный анализ. Таблицы частот, или как еще их называют одновходовые таблицы, представляют собой простейший метод анализа категориальных переменных. Таблицы частот могут быть с успехом использованы также для исследования количественных переменных, хотя при этом могут возникнуть трудности с интерпретацией результатов. Данный вид статистического исследования часто используют как одну из процедур разведочного анализа, чтобы посмотреть, каким образом различные группы наблюдений распределены в выборке, или как распределено значение признака на интервале от минимального до максимального значения. Как правило, таблицы частот графически иллюстрируются при помощи гистограмм.

Кросстабуляция (сопряжение) – процесс объединения двух (или нескольких) таблиц частот так, что каждая ячейка в построенной таблице представляется единственной комбинацией значений или уровней табулированных переменных. Кросстабуляция позволяет совместить частоты появления наблюдений на разных уровнях рассматриваемых факторов. Исследуя эти частоты, можно выявить связи между табулированными переменными и исследовать структуру этой связи. Обычно табулируются категориальные или количественные переменные с относительно небольшим числом значений. Если надо табулировать непрерывную переменную (предположим, уровень сахара в крови), то вначале ее следует перекодировать, разбив диапазон изменения на небольшое число интервалов (например, уровень: низкий, средний, высокий).

Анализ соответствий. Анализ соответствий по сравнению с частотным анализом содержит более мощные описательные и разведочные методы анализа двухвходовых и многовходовых таблиц. Метод, так же, как и таблицы сопряженности, позволяет исследовать структуру и взаимосвязь группирующих переменных, включенных в таблицу. В классическом анализе соответствий частоты в таблице сопряженности стандартизуются (нормируются) таким образом, чтобы сумма элементов во всех ячейках была равна 1. Одна из целей анализа соответствий – представление содержимого таблицы относительных частот в виде расстояний между отдельными строками и/или столбцами таблицы в пространстве более низкой размерности.

Кластерный анализ. Кластерный анализ – это метод классификационного анализа; его основное назначение – разбиение множества исследуемых объектов и признаков на однородные в некотором смысле группы, или кластеры. Это многомерный статистический метод, поэтому предполагается, что исходные данные могут быть значительного объема, т.е. существенно большим может быть как количество объектов исследования (наблюдений), так и признаков, характеризующих эти объекты. Большое достоинство кластерного анализа в том, что он дает возможность производить разбиение объектов не по одному признаку, а по ряду признаков. Кроме того, кластерный анализ в отличие от большинства математико-статистических методов не накладывает никаких ограничений на вид рассматриваемых объектов и позволяет исследовать множество исходных данных практически произвольной природы. Так как кластеры – это группы однородности, то задача кластерного анализа заключается в том, чтобы на основании признаков объектов разбить их множество на m (m – целое) кластеров так, чтобы каждый объект принадлежал только одной группе разбиения. При этом объекты, принадлежащие одному кластеру, должны быть однородными (сходными), а объекты, принадлежащие разным кластерам, – разнородными. Если объекты кластеризации представить как точки в n-мерном пространстве признаков (n – количество признаков, характеризующих объекты), то сходство между объектами определяется через понятие расстояния между точками, так как интуитивно понятно, что чем меньше расстояние между объектами, тем они более схожи.

Дискриминантный анализ. Дискриминантный анализ включает статистические методы классификации многомерных наблюдений в ситуации, когда исследователь обладает так называемыми обучающими выборками. Этот вид анализа является многомерным, так как использует несколько признаков объекта, число которых может быть сколь угодно большим. Цель дискриминантного анализ состоит в том, чтобы на основе измерения различных характеристик (признаков) объекта классифицировать его, т. е. отнести к одной из нескольких заданных групп (классов) некоторым оптимальным способом. При этом предполагается, что исходные данные наряду с признаками объектов содержат категориальную (группирующую) переменную, которая определяет принадлежность объекта к той или иной группе. Поэтому в дискриминантном анализе предусмотрена проверка непротиворечивости классификации, проведенной методом, с исходной эмпирической классификацией. Под оптимальным способом понимается либо минимум математического ожидания потерь, либо минимум вероятности ложной классификации. В общем случае задача различения (дискриминации) формулируется следующим образом. Пусть результатом наблюдения над объектом является построение k-мерного случайного вектора X = (X1, X2, …, XK), где X1, X2, …, XK – признаки объекта. Требуется установить правило, согласно которому по значениям координат вектора X объект относят к одной из возможных совокупностей i, i = 1, 2, …, n. Методы дискриминации можно условно разделить на параметрические и непараметрические. В параметрических известно, что распределение векторов признаков в каждой совокупности нормально, но нет информации о параметрах этих распределений. Непараметрические методы дискриминации не требуют знаний о точном функциональном виде распределений и позволяют решать задачи дискриминации на основе незначительной априорной информации о совокупностях, что особенно ценно для практических применений. Если выполняются условия применимости дискриминантного анализа – независимые переменные–признаки (их еще называют предикторами) должны быть измерены как минимум в интервальной шкале, их распределение должно соответствовать нормальному закону, необходимо воспользоваться классическим дискриминантным анализом, в противном случае – методом общие модели дискриминантного анализа.

Факторный анализ. Факторный анализ – один из наиболее популярных многомерных статистических методов. Если кластерный и дискриминантный методы классифицируют наблюдения, разделяя их на группы однородности, то факторный анализ классифицирует признаки (переменные), описывающие наблюдения. Поэтому главная цель факторного анализа – сокращение числа переменных на основе классификация переменных и определения структуры взаимосвязей между ними. Сокращение достигается путем выделения скрытых (латентных) общих факторов, объясняющих связи между наблюдаемыми признаками объекта, т.е. вместо исходного набора переменных появится возможность анализировать данные по выделенным факторам, число которых значительно меньше исходного числа взаимосвязанных переменных.

Деревья классификации. Деревья классификации – это метод классификационного анализа, позволяющий предсказывать принадлежность объектов к тому или иному классу в зависимости от соответствующих значений признаков, характеризующих объекты. Признаки называются независимыми переменными, а переменная, указывающая на принадлежность объектов к классам, называется зависимой. В отличие от классического дискриминантного анализа, деревья классификации способны выполнять одномерное ветвление по переменными различных типов категориальным, порядковым, интервальным. Не накладываются какие-либо ограничения на закон распределения количественных переменных. По аналогии с дискриминантным анализом метод дает возможность анализировать вклады отдельных переменных в процедуру классификации. Деревья классификации могут быть, а иногда и бывают, очень сложными. Однако использование специальных графических процедур позволяет упростить интерпретацию результатов даже для очень сложных деревьев. Возможность графического представления результатов и простота интерпретации во многом объясняют большую популярность деревьев классификации в прикладных областях, однако, наиболее важные отличительные свойства деревьев классификации – их иерархичность и широкая применимость. Структура метода такова, что пользователь имеет возможность по управляемым параметрам строить деревья произвольной сложности, добиваясь минимальных ошибок классификации. Но по сложному дереву, из-за большой совокупности решающих правил, затруднительно классифицировать новый объект. Поэтому при построении дерева классификации пользователь должен найти разумный компромисс между сложностью дерева и трудоемкостью процедуры классификации. Широкая сфера применимости деревьев классификации делает их весьма привлекательным инструментом анализа данных, но не следует полагать, что его рекомендуется использовать вместо традиционных методов классификационного анализа. Напротив, если выполнены более строгие теоретические предположения, налагаемые традиционными методами, и выборочное распределение обладает некоторыми специальными свойствами (например, соответствие распределения переменных нормальному закону), то более результативным будет использование именно традиционных методов. Однако как метод разведочного анализа или как последнее средство, когда отказывают все традиционные методы, Деревья классификации, по мнению многих исследователей, не знают себе равных.

Анализ главных компонент и классификация. На практике часто возникает задача анализа данных большой размерности. Метод анализ главных компонент и классификация позволяет решить эту задачу и служит для достижения двух целей:
– уменьшение общего числа переменных (редукция данных) для того, чтобы получить

«главные» и «некоррелирующие» переменные;
– классификация переменных и наблюдений, при помощи строящегося факторного пространства.
Метод имеет сходство с факторным анализом в постановочной части решаемых задач, но имеет ряд существенных отличий:
– при анализе главных компонент не используются итеративные методы для извлечения факторов;
– наряду с активными переменными и наблюдениями, используемыми для извлечения главных компонент, можно задать вспомогательные переменные и/или наблюдения; затем вспомогательные переменные и наблюдения проектируются на факторное пространство, вычисленное на основе активных переменных и наблюдений;
– перечисленные возможности позволяют использовать метод как мощное средство для классификации одновременно переменных и наблюдений.
Решение основной задачи метода достигается созданием векторного пространства латентных (скрытых) переменных (факторов) с размерностью меньше исходной. Исходная размерность определяется числом переменных для анализа в исходных данных.

Многомерное шкалирование. Метод можно рассматривать как альтернативу факторному анализу, в котором достигается сокращение числа переменных, путем выделения латентных (непосредственно не наблюдаемых) факторов, объясняющих связи между наблюдаемыми переменными. Цель многомерного шкалирования – поиск и интерпретация латентных переменных, дающих возможность пользователю объяснить сходства между объектами, заданными точками в исходном пространстве признаков. Показателями сходства объектов на практике могут быть расстояния или степени связи между ними. В факторном анализе сходства между переменными выражаются с помощью матрицы коэффициентов корреляций. В многомерном шкалировании в качестве исходных данных можно использовать произвольный тип матрицы сходства объектов: расстояния, корреляции и т.д. Несмотря на то, что имеется много сходства в характере исследуемых вопросов, методы многомерное шкалирование и факторный анализ имеют ряд существенных отличий. Так, факторный анализ требует, чтобы исследуемые данные подчинялись многомерному нормальному распределению, а зависимости были линейными. Многомерное шкалирование не накладывает таких ограничений, оно может быть применимо, если задана матрица попарных сходств объектов. В терминах различий получаемых результатов факторный анализ стремится извлечь больше факторов – латентных переменных по сравнению с многомерным шкалированием. Поэтому многомерное шкалирование часто приводит к проще интерпретируемым решениям. Однако более существенно то, что метод многомерное шкалирование можно применять к любым типам расстояний или сходств, в то время как факторный анализ требует, чтобы в качестве исходных данных была использована корреляционная матрица переменных или по файлу исходных данных сначала была вычислена матрица корреляций. Основное предположение многомерного шкалирования заключается в том, что существует некоторое метрическое пространство существенных базовых характеристик, которые неявно и послужили основой для полученных эмпирических данных о близости между парами объектов. Следовательно, объекты можно представить как точки в этом пространстве. Предполагают также, что более близким (по исходной матрице) объектам соответствуют меньшие расстояния в пространстве базовых характеристик. Поэтому, многомерное шкалирование – это совокупность методов анализа эмпирических данных о близости объектов, с помощью которых определяется размерность пространства существенных для данной содержательной задачи характеристик измеряемых объектов и конструируется конфигурация точек (объектов) в этом пространстве. Это пространство («многомерная шкала») аналогично обычно используемым шкалам в том смысле, что значениям существенных характеристик измеряемых объектов соответствуют определенные позиции на осях пространства. Логику многомерного шкалирования можно проиллюстрировать на следующем простом примере. Предположим, что имеется матрица попарных расстояний (т.е. сходства некоторых признаков) между некоторыми городами. Анализируя матрицу, надо расположить точки с координатами городов в двумерном пространстве (на плоскости), максимально сохранив реальные расстояния между ними. Полученное размещение точек на плоскости впоследствии можно использовать в качестве приближенной географической карты. В общем случае многомерное шкалирование позволяет таким образом расположить объекты (города в нашем примере) в пространстве некоторой небольшой размерности (в данном случае она равна двум), чтобы достаточно адекватно воспроизвести наблюдаемые расстояния между ними. В результате можно измерить эти расстояния в терминах найденных латентных переменных. Так, в нашем примере можно объяснить расстояния в терминах пары географических координат Север/Юг и Восток/Запад.

Моделирование структурными уравнениями (причинное моделирование). Наметившийся в последнее время прогресс в области многомерного статистического анализа и анализа корреляционных структур, объединенный с новейшими вычислительными алгоритмами, послужил отправной точкой для создания новой, но уже получившей признание техники моделирования структурными уравнениями (SEPATH). Эта необычайно мощная техника многомерного анализа включает методы из различных областей статистики, множественная регрессия и факторный анализ получили здесь естественное развитие и объединение.
Объектом моделирования структурными уравнениями являются сложные системы, внутренняя структура которых не известна («черный ящик»). Наблюдая параметры системы при помощи SEPATH, можно исследовать ее структуру, установить причинно-следственные взаимосвязи между элементами системы.
Постановка задачи структурного моделирования выглядит следующим образом. Пусть имеются переменные, для которых известны статистические моменты, например, матрица выборочных коэффициентов корреляции или ковариации. Такие переменные называются явными. Они могут быть характеристиками сложной системы. Реальные связи между наблюдаемыми явными переменными могут быть достаточно сложными, однако предполагаем, что имеется некоторое число скрытых переменных, которые с известной степенью точности объясняют структуру этих связей. Таким образом, с помощью латентных переменных строится модель связей между явными и неявными переменными. В некоторых задачах латентные переменные можно рассматривать как причины, а явные – как следствия, поэтому, такие модели называются причинными. Допускается, что скрытые переменные, в свою очередь, могут быть связаны между собой. Структура связей допускается достаточно сложной, однако тип ее постулируется – это связи, описываемые линейными уравнениями. Какие-то параметры линейных моделей известны, какие-то нет, и являются свободными параметрами.

Основная идея моделирования структурными уравнениями состоит в том, что можно проверить, связаны ли переменные Y и X линейной зависимостью Y = aX, анализируя их дисперсии и ковариации. Эта идея основана на простом свойстве среднего и дисперсии: если умножить каждое число на некоторую константу k, среднее значение также умножится на k, при этом стандартное отклонение умножится на модуль k. Например, рассмотрим набор из трех чисел 1, 2, 3. Эти числа имеют среднее, равное 2, и стандартное отклонение, равное 1. Если умножить все три числа на 4, то легко посчитать, что среднее значение будет равно 8, стандартное отклонение – 4, а дисперсия – 16. Таким образом, если есть наборы чисел X и Y, связанные зависимостью Y = 4X, то дисперсия Y должна быть в 16 раз больше, чем дисперсия X. Поэтому можно проверить гипотезу о том, что Y и X связаны уравнением Y = 4X, сравнением дисперсий переменных Y и X. Эта идея может быть различными способами обобщена на несколько переменных, связанных системой линейных уравнений. При этом правила преобразований становятся более громоздкими, вычисления более сложными, но основной смысл остается прежним – можно проверить, связаны ли переменные линейной зависимостью, изучая их дисперсии и ковариации.

Методы анализа выживаемости. Методы анализа выживаемости первоначально были развиты в медицинских, биологических исследованиях и страховании, но затем стали широко применяться в социальных и экономических науках, а также в промышленности в инженерных задачах (анализ надежности и времен отказов). Представьте, что изучается эффективность нового метода лечения или лекарственного препарата. Очевидно, наиболее важной и объективной характеристикой является средняя продолжительность жизни пациентов с момента поступления в клинику или средняя продолжительность ремиссии заболевания. Для описания средних времен жизни или ремиссии можно было бы использовать стандартные параметрические и непараметрические методы. Однако в анализируемых данных есть существенная особенность – могут найтись пациенты, которые в течение всего периода наблюдения выжили, а у некоторых из них заболевание все еще находится в стадии ремиссии. Также может образоваться группа больных, контакт с которыми был потерян до завершения эксперимента (например, их перевели в другие клиники). При использовании стандартных методов оценки среднего эту группу пациентов пришлось бы исключить, тем самым, потеряв с трудом собранную важную информацию. К тому же большинство этих пациентов являются выжившими (выздоровевшими) в течение того времени, которое их наблюдали, что свидетельствует в пользу нового метода лечения (лекарственного препарата). Такого рода информация, когда нет данных о наступлении интересующего нас события, называется неполной. Если есть данные о наступлении интересующего нас события, то информация называется полной. Наблюдения, которые содержат неполную информацию, называются цензурированными наблюдениями. Цензурированные наблюдения типичны, когда наблюдаемая величина представляет время до наступления некоторого критического события, а продолжительность наблюдения ограничена по времени. Использование цензурированных наблюдений составляет специфику рассматриваемого метода – анализа выживаемости. В данном методе исследуются вероятностные характеристики интервалов времени между последовательным возникновением критических событий. Такого рода исследования называются анализом длительностей до момента прекращения, которые можно определить как интервалы времени между началом наблюдения за объектом и моментом прекращения, при котором объект перестает отвечать заданным для наблюдения свойствам. Цель исследований – определение условных вероятностей, связанных с длительностями до момента прекращения. Построение таблиц времен жизни, подгонка распределения выживаемости, оценивание функции выживания с помощью процедуры Каплана – Мейера относятся к описательным методам исследования цензурированных данных. Некоторые из предложенных методов позволяют сравнивать выживаемость в двух и более группах. Наконец, анализ выживаемости содержит регрессионные модели для оценивания зависимостей между многомерными непрерывными переменными со значениями, аналогичными временам жизни. Общие модели дискриминантного анализа. Если не выполняются условия применимости дискриминантного анализа (ДА) – независимые переменные (предикторы) должны быть измерены как минимум в интервальной шкале, их распределение должно соответствовать нормальному закону, необходимо воспользоваться методом общие модели дискриминантного анализа (ОДА). Метод имеет такое название, потому что в нем для анализа дискриминантных функций используется общая линейная модель (GLM). В этом модуле анализ дискриминантных функций рассматривается как общая многомерная линейная модель, в которой категориальная зависимая переменная (отклик) представляется векторами с кодами, обозначающими различные группы для каждого наблюдения. Метод ОДА имеет ряд существенных преимуществ перед классическим дискриминантным анализом. Например, не устанавливается никаких ограничений на тип используемого предиктора (категориальный или непрерывный) или на тип определяемой модели, возможен пошаговый выбор предикторов и выбор наилучшего подмножества предикторов, в случае наличия в файле данных кросс-проверочной выборки выбор наилучшего подмножества предикторов можно провести на основе долей ошибочной классификации для кросс-проверочной выборки и т.д.

Временные ряды. Временные ряды – это наиболее интенсивно развивающееся, перспективное направление математической статистики. Под временным (динамическим) рядом подразумевается последовательность наблюдений некоторого признака X (случайной величины) в последовательные равноотстоящие моменты t. Отдельные наблюдения называются уровнями ряда и обозначаются xt, t = 1, …, n. При исследовании временного ряда выделяются несколько составляющих:
$x_t=u_t+y_t+c_t+e_t$, t = 1, …, n,
где u_t – тренд, плавно меняющаяся компонента, описывающая чистое влияние долговременных факторов (убыль населения, уменьшение доходов и т.д.); – сезонная компонента, отражающая повторяемость процессов в течение не очень длительного периода (дня, недели, месяца и т.д.); ct – циклическая компонента, отражающая повторяемость процессов в течение длительных периодов времени свыше одного года; t – случайная компонента, отражающая влияние не поддающихся учету и регистрации случайных факторов. Первые три компоненты представляют собой детерминированные составляющие. Случайная составляющая образована в результате суперпозиции большого числа внешних факторов, оказывающих каждый в отдельности незначительное влияние на изменение значений признака X. Анализ и исследование временного ряда позволяют строить модели для прогнозирования значений признака X на будущее время, если известна последовательность наблюдений в прошлом.

Нейронные сети. Нейронные сети представляют собой вычислительную систему, архитектура которой имеет аналогию с построением нервной ткани из нейронов. На нейроны самого нижнего слоя подаются значения входных параметров, на основании которых нужно принимать определенные решения. Например, в соответствии со значениями клинико-лабораторных показателей больного надо отнести его к той или иной группе по степени тяжести заболевания. Эти значения воспринимаются сетью как сигналы, передающиеся в следующий слой, ослабляясь или усиливаясь в зависимости от числовых значений (весов), приписываемых межнейронным связям. В результате на выходе нейрона верхнего слоя вырабатывается некоторое значение, которое рассматривается как ответ – отклик всей сети на входные параметры. Для того, чтобы сеть работала ее надо «натренировать» (обучить) на данных для которых известны значения входных параметров и правильные отклики на них. Обучение состоит в подборе весов межнейронных связей, обеспечивающих наибольшую близость ответов к известным правильным ответам. Нейронные сети могут быть использованы для классификации наблюдений.

Планирование экспериментов. Искусство располагать наблюдения в определенном порядке или проводить специально спланированные проверки с целью полного использования возможностей этих методов и составляет содержание предмета «планирование эксперимента». В настоящее время экспериментальные методы широко используются как в науке, так и в различных областях практической деятельности. Обычно основная цель научного исследования состоит в том, чтобы показать статистическую значимость эффекта воздействия определенного фактора на изучаемую зависимую переменную. Как правило, основная цель планирования экспериментов заключается в извлечении максимального количества объективной информации о влиянии изучаемых факторов на интересующий исследователя показатель (зависимую переменную) с помощью наименьшего числа дорогостоящих наблюдений. К сожалению, на практике, в большинстве случаев, недостаточное внимание уделяется планированию исследований. Собирают данные (столько, сколько могут собрать), а потом уже проводят статистическую обработку и анализ. Но сам по себе правильно проведенный статистический анализ недостаточен для достижения научной достоверности, поскольку качество любой информации, получаемой в результате анализа данных, зависит от качества самих данных. Поэтому планирование экспериментов находит все большее применение в прикладных исследованиях. Целью методов планирования экспериментов является изучение влияния определенных факторов на исследуемый процесс и поиск оптимальных уровней факторов, определяющих требуемый уровень течения данного процесса.

Карты контроля качества. В условиях современного мира чрезвычайно актуальным является проблема качества не только выпускаемой продукции, но и услуг оказываемых населению. От успешного решения этой важной проблемы в значительной степени зависит благополучие любой фирмы, организации или учреждения. Качество продукции и услуг формируется в процессе научных исследований, конструкторских и технологических разработок, обеспечивается хорошей организацией производства и услуг. Но изготовление продукции и оказание услуг независимо от их вида всегда связано с определенным непостоянством условий производства и предоставления. Это приводит к некоторой вариабельности признаков их качества. Поэтому, актуальными являются вопросы разработки методов контроля качества, которые позволят своевременно выявить признаки нарушения технологического процесса или оказания услуг. При этом, для достижения и поддержания высокого уровня качества, удовлетворяющего потребителя нужны методы, направленные не на устранение дефектов готовой продукции и несоответствий услуг, а на предупреждение и прогнозирование причин их появления. Контрольная карта – это инструмент, позволяющий отслеживать ход протекания процесса и воздействовать на него (с помощью соответствующей обратной связи), предупреждая его отклонения от предъявленных к процессу требований. Инструментарий карт контроля качества широко использует статистические методы, основанные на теории вероятностей и математической статистики. Применение статистических методов позволяет при ограниченных объемах анализируемых изделий с заданной степенью точности и достоверности судить о состоянии качества выпускаемой продукции. Обеспечивает прогнозирование, оптимальное регулирование проблем в области качества, принятие верных управленческих решений не на основе интуиции, а при помощи научного изучения и выявления закономерностей в накапливаемых массивах числовой информации.

1.5 Лекция № 5
Тема: «Задачи теории статистических решений»

1.5.1 Вопросы лекции:

1. Метод внутренней точки, метод внешней точки.

2. Комбинированный метод. оптимизационные методы получения детерминированных оценок (методы линейного программирования, квадратичного программирования, выпуклого программирования, динамическое программирование, принцип максимума, оптимизация в функциональных пространствах

1.5.2. Краткое содержание вопросов:

1. Метод внутренней точки, метод внешней точки.

Методы штрафных функций относятся к группе непрямых методов решения задач нелинейного программирования:

$f(x) \to \min;$

$g_i(x) = 0,\ i = 1, ..., k;$

$h_j(x) \leq 0,\ j = 1, ..., m;$

$a \leq x \leq b.$

Они преобразуют задачу с ограничениями в последовательность задач безусловной оптимизации некоторых вспомогательных функций. Последние получаются путем модификации целевой функции с помощью функций-ограничений таким образом, чтобы ограничения в явном виде в задаче оптимизации не фигурировали. Это обеспечивает возможность применения методов безусловной оптимизации. В общем случае вспомогательная функция имеет вид

$F(x,a) = f(x) + \Phi(x, a).$

Здесь *f(x)* - целевая функция задачи оптимизации; *Φ(x, a)* - “штрафная” функция; параметр *a* $>$0. Точку безусловного минимума функции *F(x, a)* будем обозначать через *x(a)*.

В зависимости от вида *Φ(x, a)* различают методы внутренних штрафных, или барьерных, функций и методы внешних штрафных функций.

Методы внутренних штрафных функций

Эти методы применяются для решения задач нелинейного программирования с ограничениями-неравенствами. В рассматриваемых методах функции *Φ(x, a)* подбирают такими, чтобы их значения неограниченно возрастали при приближении к границе допустимой области *G* (Рис. 3.2). Иными словами, приближение к границе “штрафуется” резким увеличением значения функции *F(x, a)*. На границе *G* построен “барьер”, препятствующий нарушению ограничении в процессе безусловной минимизации *F(x, a)*. Поиск минимума вспомогательной функции *F(x, a)* необходимо начинать с внутренней точки области *G* . При этом в процессе оптимизации траектория спуска никогда не выйдет за пределы допустимой области. Все перечисленные особенности функции Φ *(x, a)* определили наименование рассматриваемой группы методов.

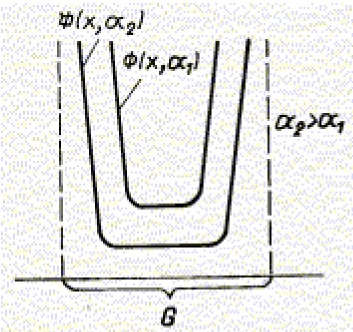


Рис. 3.2. Внутренняя штрафная функция

Таким образом, внутренняя штрафная функция *Φ(x, a)* может быть определена следующим образом:

$$\Phi(x,a) = \begin{cases} \to \infty, \text{если } x \notin G, \\ \to 0, \text{если } x \in G, x \notin dG, a \to 0; \\ \to \infty, \text{если } x \in G, x \to dG. \end{cases}$$

Здесь *dG* -граница области *G*.

Общий вид внутренней штрафной функции

$$\Phi(x,a) = a\left(\sum_{j=1}^{m} \varphi_j(h_j(x))\right),$$

где φ_j - непрерывные дифференцируемые функции, определяемые ограничениями-неравенствами исходной задачи нелинейного программирования. Вспомогательная функция *F(x, a)* при этом имеет форму

$$F(x,a) = f(x) + a\sum_{j=1}^{m} \varphi_j(h_j(x)).$$

Она определена в области *G* и неограниченно возрастает, если $h_j(x)$ -> 0 для некоторого *j*. В качестве внутренних штрафных функций используют, например, такие:

$$\Phi(x,a) = a\sum_{j=1}^{m}\frac{1}{h_j(x)};\quad \Phi(x,a) = -a\sum_{j=1}^{m}\ln h_j(x).$$

Алгоритм метода внутренних штрафных функций состоит в следующем. В качестве начальной точки $x[0]$ выбирается произвольная внутренняя точка области G. Задается некоторая монотонно убывающая сходящаяся к нулю последовательность $\{a_k\}$, $k = 1, 2, \ldots$, положительных чисел. Для первого элемента a_1 этой последовательности решается задача безусловной минимизации функции $F(x, a)$, в результате чего определяется точка $x(a_1)$. Эта точка используется в качестве начальной для решения задачи поиска минимума функции $F(x, a_2)$, где $a_2 < a_1$, и т. д. Таким образом, решается последовательность задач безусловной минимизации функций $F(x, a_k)$, $k = 1, 2, \ldots$, причем решение предыдущей задачи $x(a_k)$ используется в качестве начальной точки для поиска последующего вектора $x(a_{k+1})$. Последовательность полученных таким образом точек $x(a_k)$ сходится к оптимальному решению исходной задачи - локальному минимуму x^*. Вычисления прекращают при выполнении условий:

$|f(x[k]) - f(x[k-1])| \leq$ □;

$\|x[k] - x[k-1]\| \leq$ □;

Здесь □, □ - заданные числа, определяющие точность вычислений.

Можно показать, что рассмотренный метод внутренних штрафных функций обладает следующими свойствами:

1) $\lim\limits_{k\to\infty} a_k \sum_{j=1}^{m} \varphi_j(h_j(x[k])) = 0$;

2) $\lim\limits_{k\to\infty} f(x[k]) = f(x^*)$ и $f(x[k])$ монотонно убывает;

3) $\lim\limits_{k\to\infty} F(x[k]a_k) = f(x^*)$

Эти свойства справедливы для задач, содержащих непрерывные функции и имеющих локальные минимумы внутри области G.

Методы внешних штрафных функций

Данные методы применяются для решения задачи оптимизации в общей постановке, т. е. при наличии как ограничений-неравенств, так и ограничений-равенств. В рассматриваемых методах функции $\Phi(x, a)$ выбирают такими, что их значения равны нулю внутри и на границе допустимой области G, а вне ее -положительны и возрастают тем больше, чем сильнее нарушаются ограничения (Рис. 3.3). Таким образом, здесь “штрафуется” удаление от допустимой области G.

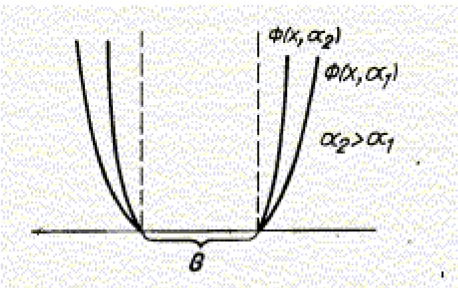


Рис. 3.3. Внешняя штрафная функция

Внешняя штрафная функция Ф(x, a) в общем случае может быть определена следующим образом:

$$\Phi(x,a)=\begin{cases} 0,\text{`если`}\,x \in G; \\ \to \infty,\text{`если`}\,x \in G,\text{`}a \to \infty. \end{cases}$$

Поиск минимума вспомогательной функции *F(x, a)* можно начинать из произвольной точки. В большинстве случаев она является недопустимой, поэтому траектория спуска располагается частично вне допустимой области. Если минимум целевой функции расположен на границе допустимой области, то эта траектория полностью находится снаружи области G. Перечисленные особенности функции *Φ(x, a)* определили название данной группы методов. Общий вид внешней штрафной функции:

$$\Phi(x,a)=a\left(\sum_{i=1}^{k}\psi_i(g_i(x))+\sum_{j=1}^{m}\varphi_j(h_j(x))\right),$$

где□□$_j$, □ $_j$ - функции, определяемые соответственно ограничениями-равенствами и неравенствами исходной задачи нелинейного программирования. Вспомогательная функция *F(x, a)* при этом имеет форму

$$F(x,a)=f(a)+a\left(\sum_{i=1}^{k}\psi_i(g_i(x))+\sum_{j=1}^{m}\varphi_j(h_j(x))\right)$$

Одна из применяемых внешних штрафных функций имеет вид

$$\Phi(x,a)=a\left(\sum_{i=1}^{k}(\max(0,g_i(x)))^2+\sum_{j=1}^{m}(h_j(x))^2\right)$$

Здесь $\max(0,g_i(x))=\begin{cases} 0,\text{если} & g_i(x)\le 0; \\ g_j(x),\text{если} & g_i(x)>0 \end{cases}$

Алгоритм метода внешних штрафных функций формулируется так же, как и алгоритм метода внутренних штрафных функций, и обладает аналогичными свойствами. Однако в этом случае не требуется, чтобы начальная точка $x[0] \in G$, а последовательность $\{a_k\}$, k = 1, 2, ..., положительных чисел должна быть монотонно возрастающей.

Анализ методов штрафных функций позволяет сделать следующие выводы об их вычислительных свойствах. В соответствии с методами внутренних штрафных функций ведут поиск решения, не выходя за пределы допустимой области. Это весьма важно в тех случаях, когда целевая функция или ограничения не определены за пределами допустимого множества. Кроме того, прервав вычисления в любой момент времени, мы всегда получим допустимое решение. Однако для задания в качестве начальной некоторой допустимой точки иногда требуется решать задачу, по сложности сравнимую с исходной задачей нелинейного программирования. В этом смысле метод внешних штрафных функций предпочтительнее, так как он обеспечивает решение из любой начальной точки. В результате программирование для ЭВМ алгоритмов внешних штрафных функций существенно упрощается. Общим недостатком методов штрафных функций является сложность вспомогательной функции *F(x, a),* которая часто имеет овражную структуру. Степень овражности увеличивается с увеличением *a.* Кроме того, при больших значениях *a* точность вычислений минимума *F(x, a)* сильно уменьшается из-за ошибок округления ЭВМ.

2. Комбинированные алгоритмы штрафных функций

Для задач нелинейного программирования с ограничениями-равенствами методы внутренних штрафных функций неприменимы. Однако при практических расчетах в ряде случаев необходимо выполнение некоторых ограничений-неравенств в течение всего процесса оптимизации. В подобных обстоятельствах используют *комбинированные алгоритмы,* учитывающие особенности внутренних и внешних штрафных функций. Вспомогательная функция *F(x, a)* включает в себя слагаемые в виде внутренних штрафных функций □ *(x, a)* для ограничений-неравенств и внешних штрафных функций □□*(x, a)* для ограничений-равенств:

F(x, a) = f(x) + Φ(x, a) + □□*(x, a)* .

В таком случае неравенства будут выполнены на протяжении всего вычислительного процесса, а равенства - при приближении к минимуму. Например, в качестве внутренней можно использовать логарифмическую штрафную функцию, а в качестве внешней - квадратичную функцию, т. е.

$$F(x,a_k) = f(x) + a_k \sum_{j=1}^{m} \ln(-h(x)) + \frac{1}{a_k} \sum_{i=1}^{k} g_i^2(x)$$
.

В качестве начальной точки выбирается вектор $x[0]$, удовлетворяющий условиям $h_j(x) > 0$, $j = 1, \ldots, m$. Последовательность $\{a_k\}$, $k = 1, 2, \ldots$, положительных чисел является монотонно убывающей и сходящейся к нулю.

Выбор начальной точки допустимой области

В данном случае задача состоит в поиске точки, удовлетворяющей строгим неравенствам $h_i(x) > 0$, $j = 1, \ldots, m$. Рассмотрим два множества индексов:

$J_1 = \{j \mid h_j(\bar{x}) < 0, j = 1, \ldots, m\}$

$J_2 = \{j \mid h_j(\bar{x}) \geq 0, j = 1, \ldots, m\}$

Поиск внутренней точки должен состоять в том, чтобы перейти от точки $(\bar{x})$ к новой точке $(\bar{x})$, в которой удовлетворяются некоторые из ограничений множества J_2. При этом ни одно из условий множества J_1 не должно нарушаться. Затем необходимо перейти от точки $(\bar{x})$ к другой точке, и так далее до тех пор, пока не будут удовлетворены все ограничения. Эта процедура реализуется минимизацией функции

$$W(x,a_k) = \sum_{j \in J_2} h_j(x) + a_k V(x)$$
,

где $V(x)$ - внутренняя штрафная функция одного из видов, представленных выше, относительно ограничений из множества J_1. Последовательность $\{a_k\}$, $k = 1, 2, \ldots$, сходится к нулю. В процессе минимизации производится проверка знаков функций h_j, $j \in J_2$. Как только какое-либо из этих ограничений удовлетворяется, оно переводится во второе слагаемое, т.е. формируется соответствующая штрафная функция $V(x)$. Минимизация полученной в результате новой функции *W(x, a.)* осуществляется из последней найденной точки $x[k]$.

Оптимизационные методы получения детерминированных оценок

2.1. Возникновение задач оптимизации

Людям свойственно стремление к лучшему, и если им приходится выбирать из нескольких возможностей, то желание найти среди них оптимальную представляется вполне естественным.

Слово «оптимальный» происходит от латинского optimus, что значит – наилучший, совершенный. Для того, чтобы найти оптимальную из возможностей, приходится решать задачи на отыскание максимума или минимума, то есть наибольших и наименьших значений каких-либо величин. Оба эти понятия – максимум и минимум – объединяются единым термином «экстремум» (от латинского «extremum», означающего «крайнее»). Поэтому задачи на отыскание максимума или минимума называют экстремальными задачами.

Великий русский математик П.Л.Чебышев (1821-1894) писал, что особенную важность имеют те методы науки, которые позволяют решать задачу, общую для всей практической деятельности человека: как располагать своими средствами для достижения по возможности большей выгоды. С такими задачами приходится иметь дело представителям самых разных специальностей – инженеры-технологи стремятся так организовать производство, чтобы на имеющемся станочном парке сделать как можно больше продукции, конструкторы ломают голову, стремясь сделать наилегчайшим прибор на космическом корабле, экономисты стараются так спланировать прикрепление заводов к источникам сырья, чтобы транспортные расходы оказались наименьшими.

Но не только людям приходится решать подобные задачи. Бессознательно с ними справляются и некоторые виды насекомых и других живых существ. Например, форма ячеек пчелиных сот такова, что при заданном объеме на них идет наименьшее количество воска. Неумолимый естественный отбор привел к тому, что выжили лишь пчелы, тратившие меньше всего усилий на строительство сот.

Пример 1.1. Предположим, что надо вытесать из цилиндрического бревна брус наибольшего объема. Эта задача сводится к тому, чтобы вписать в круг данного радиуса R прямоугольник наибольшей площади (рис. 2.1).

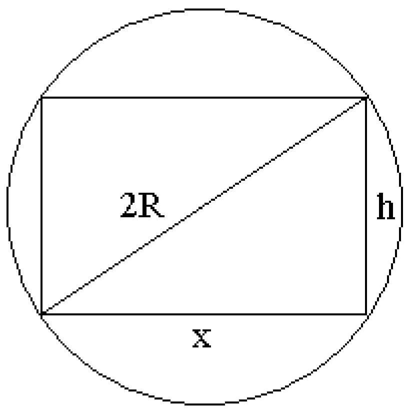


Рис. 2.1. Прямоугольник, вписанный в круг

Если обозначить ширину прямоугольника через x, то его высота h равна $\sqrt{4R^2 - x^2}$, а площадь выражается следующей формулой $S = x\sqrt{4R^2 - x^2}$. И теперь нам надо найти x, при котором функция S принимает наибольшее значение. Вычислим производную и приравняем ее нулю, тогда получим

$$S' = \sqrt{4R^2 - x^2} - x\frac{2x}{2\sqrt{4R^2 - x^2}} = 0, \qquad 0 \le x \le 2R.$$

Отсюда $x = R\sqrt{2}$. При $x = 0$ и $x = 2R$ площадь $S = 0$, а при $x = R\sqrt{2}$ площадь $S = 2R^2$. Это значит, что $x = R\sqrt{2}$ – точка максимума функции S, и прямоугольник должен быть квадратом.

Пример 1.2. По каналу, имеющему форму, приведенную на рис. 2.2, производится сплав бревен. Определить максимальную длину бревна, которое может пройти поворот этого канала. Толщиной бревна следует пренебречь.

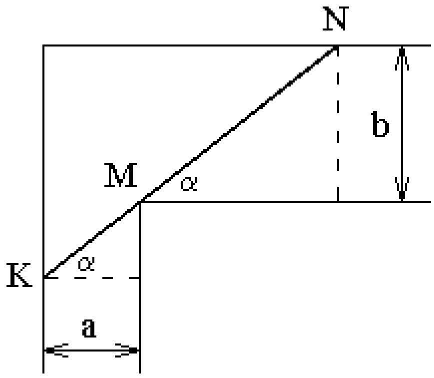


Рис. 2.2. Форма канала

Длина бревна, концы которого опираются на внешние части канала, равна $L = |KM| + |MN|$, или $L = \dfrac{a}{\cos\alpha} + \dfrac{b}{\sin\alpha}$. Угол α может изменяться в переделах от 0 до $\dfrac{\pi}{2}$.

Если $\alpha = 0$, или $\alpha = \dfrac{\pi}{2}$, то значение L будет бесконечным, и бревно не пройдет по каналу. Бревно с концами, опирающимися на внешние части канала, пройдет по нему только тогда, когда значение L будет наименьшим. Вычислим производную от функции L и приравняем ее нулю. Получим

$$L' = -\frac{a\sin\alpha}{\cos^2\alpha} + \frac{b\cos\alpha}{\sin^2\alpha} = 0,$$

откуда

$$tg^3\alpha = \frac{b}{a}, \text{ или } tg\alpha = \sqrt[3]{\frac{b}{a}}.$$

Так как $\cos\alpha = \dfrac{1}{\sqrt{1+tg^2\alpha}}$ и $\sin\alpha = \dfrac{tg\alpha}{\sqrt{1+tg^2\alpha}}$, то максимальная длина бревна, которое пройдет поворот, равна

$$L = a\sqrt{1+tg^2\alpha} + b\frac{\sqrt{1+tg^2\alpha}}{tg\alpha} = \left(a + \frac{b}{tg\alpha}\right)\sqrt{1+tg^2\alpha},$$

или

$$L = \left(a + b\sqrt[3]{\frac{a}{b}}\right)\sqrt{1+\sqrt[3]{\frac{b^2}{a^2}}} = \left(a^{\frac{2}{3}} + b^{\frac{2}{3}}\right)^{\frac{3}{2}}.$$

Пример 1.3. Балка наибольшей прочности. Основным элементом любой строительной конструкции является балка. Прочность балки зависит от того, какую форму имеет ее поперечное сечение. При этом высота сечения играет значительно большую роль, чем ее ширина. В сопротивлении материалов доказывается, что прочность балки с прямоугольным сечением пропорциональна ширине балки b и квадрату ее высоты h. Иными словами, прочность такой балки равна kbh^2, где k – коэффициент пропорциональности, зависящий от длины балки, материала, из которого она сделана, и т.д.

Деревянные балки обычно приходится вытесывать из круглых бревен. В связи с этим возникает следующая задача: из бревна радиуса R сделать балку наибольшей прочности.

На рис. 2.1 изображено поперечное сечение бревна. Выразим прочность балки как функцию от ее ширины x. Так как высота балки $h = \sqrt{4R^2 - x^2}$, то ее прочность равна $y = kx(4R^2 - x^2)$. При этом x изменяется от 0 до $2R$.

Функция $y = kx(4R^2 - x^2)$ обращается в нуль при $x = 0$ и $x = 2R$ и положительна между этими значениями. Значит, она имеет максимум, лежащий между 0 и $2R$. Производная этой функции $y' = k(4R^2 - 3x^2)$ обращается в нуль при $x = \dfrac{2R\sqrt{3}}{3}$. Это и есть оптимальное значение ширины балки. Высота балки такой ширины равна $h = \dfrac{2R\sqrt{6}}{3}$, и отношение $\dfrac{h}{x}$ равно $\sqrt{2} \approx 1{,}4 = \dfrac{7}{5}$. Именно такое отношение высоты вытесываемой балки к ее ширине предписывается правилами производства строительных работ.

Пример 1.4. Законы отражения и преломления света. В основе этих законов лежит общий принцип, который формулируется следующим образом: луч света, попадающий из точки A в точку B, движется по такому пути, что время, затраченное им на этот путь, минимально (рис. 2.3).

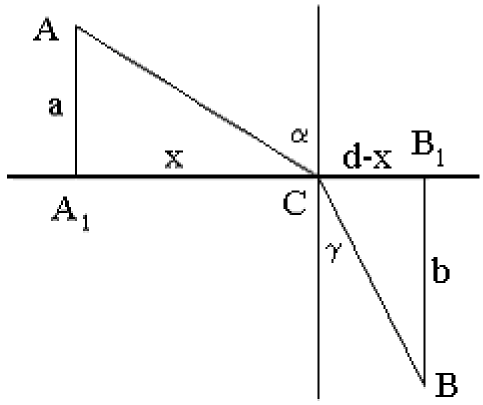


Рис. 2.3. К закону преломления света

Выведем из этого принципа закон преломления света при переходе из среды, где свет распространяется со скоростью v_1 в среду, где он распространяется со скоростью v_2. В этом случае время, затраченное светом на путь из A в B, выражается формулой

$$t = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (d - x)^2}}{v_2},$$

где d – расстояние между проекциями точек A и B на линию раздела двух сред, a и b – расстояния самих точек от этой линии, x – расстояние между проекцией A_1 точки A и точки преломления C.

Дифференцируя эту функцию и приравнивая производную нулю, получим

$$\frac{x}{v_1\sqrt{a^2 + x^2}} = \frac{d - x}{v_2\sqrt{b^2 + (d - x)^2}},$$

то есть

$$\frac{\sin\alpha}{v_1} = \frac{\sin\gamma}{v_2}.$$

Это и есть закон преломления света.

В рассмотренных задачах исследуемая функция зависела от одного аргумента x. Однако, часто необходимо находить наибольшее и наименьшее значение функций, зависящих от многих переменных. Инженер-химик должен искать такие значения температуры, давления, состава смеси и т.д., при которых данный химический процесс пойдет как можно быстрее. А агроном должен так определять сроки посева, сбора урожая, количества удобрений, чтобы при заданных затратах и заданном количестве рабочей силы и техники собрать как можно больший урожай. Иногда количество переменных, от которых зависит ответ, достигает миллионов – так обстоит дело, например, в задачах экономики, где надо составить оптимальный план, оперируя выпуском продукции на многих тысячах предприятий и устанавливая различные варианты распределения этой продукции.

1.2. Постановка задачи оптимального проектирования

В общем виде *задача оптимального проектирования* ставится следующим образом: требуется найти совокупность чисел $x_1, x_2, ..., x_n$, для которых функция

$$z = f(x_1, x_2, ..., x_n) \qquad (2.1)$$

достигает наибольшего (наименьшего) значения, и при этом выполняются условия

$$\begin{cases} g_1(x_1, x_2, ..., x_n) \le b_1 \\ g_2(x_1, x_2, ..., x_n) \le b_2 \\ ... \\ g_m(x_1, x_2, ..., x_n) \le b_m \end{cases} . \qquad (2.2)$$

Функция (2.1) носит название *целевой функции* или функции цели, неравенства (2.2) образуют *систему ограничений*. В этой системе могут быть также неравенства вида $\ge$, или равенства.

По смыслу задачи неизвестные $x_1, x_2, ..., x_n$, как правило, являются неотрицательными, то есть $x_j \ge 0$, $j = 1,2,...,n$. Эти условия могут содержаться среди неравенств (2.2), но могут также быть выписаны отдельно. Часть или даже все неизвестные задачи иногда должны быть целыми, тогда эти условия также включаются в систему ограничений.

Количество ограничений m и число неизвестных n характеризуют размерность задачи оптимизации.

Любой набор переменных, удовлетворяющих системе ограничений, называется *допустимым решением* или *планом*. Совокупность всех допустимых решений называется допустимым множеством. Численные значения целевой функции позволяют определить качество различных допустимых решений в соответствии с выбранным критерием. *Оптимальное решение* (*оптимальный план*) представляет собой такое допустимое решение, при котором значение целевой функции достигает экстремальной величины.

В теории принятия решений часто возникают проблемы, математические модели которых представляют собой модели оптимизации (задачи математического программирования). В состав этих моделей входят:

- целевая функция, выражающая в математической форме поставленную цель с точки зрения выбранного критерия оптимальности;
- система ограничений, то есть соотношения, которым должно удовлетворять решение данной задачи.

В зависимости от вида целевой функции и ограничительных условий в задачах оптимизации принято выделять следующие разделы:

- *линейное программирование*, в котором целевая функция, а также уравнения и неравенства системы ограничений линейны;
- *квадратичное программирование*, в котором целевая функция квадратична и выпукла, а допустимое множество определяется линейными равенствами и неравенствами;
- *выпуклое программирование*, в котором целевая функция и допустимое множество выпуклы;
- *дискретное программирование*, в котором допустимое множество дискретно, например, состоит из точек с целочисленными координатами;
- *сепарабельное программирование*, в котором целевая функция и ограничения являются сепарабельными функциями, т.е. представляют собой сумму функций, каждая из которых зависит только от одной переменной;
- *динамическое программирование*, в котором процесс оптимизации разбивается на ряд последовательных этапов;
- *стохастическое программирование*, в котором информация о задаче оптимизации носит элементы неопределенности, и некоторые ее параметры являются случайными величинами.

Основным и важнейшим методом линейной оптимизации является в настоящее время *симплексный* метод или метод *последовательного улучшения базисного плана.* Метод был разработан Дж.Данцигом в 1949 году. Но еще раньше, в 1939 году, советским ученым академиком Л.В.Канторовичем для решения задач линейного программирования был предложен так называемый метод разрешающих множителей, незначительно отличающийся от симплексного метода. Симплекс-метод дает возможность решать задачи линейного программирования как вручную, так и на вычислительных машинах. Через конечное число шагов (симплексных таблиц) или получается оптимальное решение или обнаруживается неразрешимость задачи линейного программирования.

Для решения общей задачи нелинейной оптимизации существует довольно много алгоритмов, однако лишь немногие оказываются эффективными для задач большой размерности. Ни один из этих алгоритмов не имеет по отношению к другим таких преимуществ, чтобы его можно было считать универсальным средством решения любых задач нелинейного программирования. При сравнении алгоритмов следует использовать следующие критерии: надежность, скорость решения, время подготовки задачи для решения, точность решения, степень выполнения ограничивающих условий. Методы нелинейной оптимизации принято классифицировать в зависимости от порядка производных, которые используются для максимизации (минимизации) целевой функции:

- методы нулевого порядка (методы поиска), при которых для поиска точки экстремума используются только значения целевой функции;
- методы первого порядка, при которых используются значения целевой функции и ее первых частных производных;
- методы второго порядка, при которых используются значения целевой функции и ее первых и вторых частных производных;

К *методам поиска* относятся: метод покоординатного спуска Пауэлла, метод Хука-Дживса, метод Розенброка, метод деформируемого многогранника (симплексный метод Нелдера и Мида) и его модификация в виде комплексного метода Бокса для нелинейной оптимизации с ограничениями, методы случайного поиска.

К *методам 1-го порядка* относятся градиентные методы, метод сопряженных направлений, метод переменной метрики (Дэвидона-Флетчера-Пауэлла).

К *методам 2-го порядка* относится метод Ньютона.

Характерной особенностью вычислительной стороны методов решения задач оптимизации является то, что практическое использование этих методов требует огромной вычислительной работы, которую без ЭВМ реализовать крайне трудно, а в ряде случаев – невозможно. В первую очередь это связано с тем, что задачи оптимизации, формализующие реальные производственные ситуации, являются задачами большой размерности, недоступными для ручного счета.

Практическую реализацию методов оптимизации для учебных задач невысокой размерности удобно проводить средствами табличного процессора Microsoft Excel. Вычислительные возможности оптимизации объединены здесь с большим набором функций, присущих текстовому и графическому редакторам и другим приложениям пакета Microsoft Office. Excel позволяет выполнять линейную и нелинейную оптимизацию (для достаточно гладких функций f и g_i, входящих в задачу), осуществлять прогнозирование и поддержку принятия решений.

Важное достоинство табличного процессора состоит в возможности автоматического пересчета всех данных, связанных функциональными зависимостями, при изменении любого компонента таблицы. Тем самым студент может в определенной степени управлять процессом оптимизации и принятия решений.

1.6 Лекция № 6
Тема: «Принятие решения на основе экспертных оценок»

1.6.1 Вопросы лекции:

1. Организация проведения экспертизы.

2. Особенности проведения подготовительного этапа, этапа функционирования рабочей группы, заключительного этапа.

1.6.2. Краткое содержание вопросов:

1. Организация проведения экспертизы

Примеры методов экспертных оценок. Как будет изменяться экономическая обстановка с течением времени? Что будет с окружающей природной средой через десять лет? Как изменится экологическая обстановка? Будет ли обеспечена экологическая безопасность промышленных производств или же вокруг станет простираться рукотворная пустыня? Достаточно вдуматься в эти постановки естественных вопросов, проанализировать, как десять или тем более двадцать лет назад мы представляли себе сегодняшний день, чтобы понять, что стопроцентно надежных прогнозов просто не может быть. Вместо утверждений с конкретными числами можно ожидать лишь качественных оценок. Тем не менее мы, менеджеры, экономисты, инженеры, должны принимать решения, например, об экологических и иных проектах и инвестициях, последствия которых скажутся через десять, двадцать и т.д. лет. Как быть? Остается обратиться к методам экспертных оценок. Что это за методы?

Бесспорно совершенно, что для принятия обоснованных решений необходимо опираться на опыт, знания и интуицию специалистов. После второй мировой войны в рамках кибернетики, теории управления, менеджмента и исследования операций стала развиваться самостоятельная дисциплина - теория и практика экспертных оценок.

Методы экспертных оценок - это методы организации работы со специалистами-экспертами и обработки мнений экспертов. Эти мнения обычно выражены частично в количественной, частично в качественной форме. Экспертные исследования проводят с целью подготовки информации для принятия решений ЛПР (напомним, ЛПР – лицо принимающее решение). Для проведения работы по методу экспертных оценок создают Рабочую группу (сокращенно РГ), которая и организует по поручению ЛПР деятельность экспертов, объединенных (формально или по существу) в экспертную комиссию (ЭК).

Экспертные оценки бывают *индивидуальные* и *коллективные*. *Индивидуальные оценки* - это оценки одного специалиста. Например, преподаватель единолично ставит отметку студенту, а врач - диагноз больному. Но в сложных случаях заболевания или угрозе отчисления студента за плохую учебу обращаются к *коллективному* мнению - симпозиуму врачей или комиссии преподавателей. Аналогичная ситуация - в армии. Обычно командующий принимает решение единолично. Но в сложных и ответственных ситуациях проводят военный совет. Один из наиболее известных примеров такого рода - военный совет 1812 г. в Филях, на котором под председательством М.И. Кутузова решался вопрос: "Давать или не давать французам сражение под Москвой?"

Другой простейший пример экспертных оценок - оценка номеров в КВН. Каждый из членов жюри поднимают фанерку со своей оценкой, а технический работник вычисляет среднюю арифметическую оценку, которая и объявляется как коллективное мнение жюри (ниже увидим, что такой подход некорректен с точки зрения теории измерений).

В фигурном катании процедура усложняется - перед усреднением *отбрасываются самая большая и самая маленькая оценки.* Это делается для того, чтобы не было соблазна завысить оценку одной спортсменке (например, соотечественнице) или занизить другой. Такие резко выделяющиеся из общего ряда оценки будут сразу отброшены.

Экспертные оценки часто используются при выборе, например:

- одного варианта технического устройства для запуска в серию из нескольких образцов,

- группы космонавтов из многих претендентов,

- набора проектов научно-исследовательских работ для финансирования из массы заявок,

- получателей экологических кредитов из многих желающих,

- при выборе инвестиционных проектов для реализации среди представленных, и т.д.

Существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, учатся друг у друга, и неверные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем их усреднения позволяли принимать обоснованные решения. В других - число экспертов растет в процессе проведения экспертизы, например, при использовании метода "снежного кома" (о нем - дальше).

Не меньше существует и методов обработки ответов экспертов, в том числе весьма насыщенных математикой и компьютеризированных. Многие из них основаны на достижениях статистики объектов нечисловой природы и других современных методах прикладной статистики.

Один из наиболее известных методов экспертных оценок - это *метод "Дельфи"*. Название дано по ассоциации с древним обычаем для получения поддержки при принятии решений обращаться в Дельфийский храм. Он был расположен у выхода ядовитых вулканических газов. Жрицы храма, надышавшись отравы, начинали пророчествовать, произнося непонятные слова. Специальные "переводчики" - жрецы храма толковали эти слова и отмечали на вопросы пришедших со своими проблемами паломников. По традиции говорят, что Дельфийский храм находился в Греции. Но там нет вулканов. Видимо, он был в Италии - у Везувия или Этны, а сами описанные предсказания происходили в XII-XIV вв. Это вытекает из высшего достижения современной исторической науки - новой статистической хронологии.

В США в 1960-х годах методом Дельфи назвали экспертную процедуру прогнозирования научно-технического развития. В первом туре эксперты называли вероятные даты тех или иных будущих свершений. Во втором туре каждый эксперт знакомился с прогнозами всех остальных. Если его прогноз сильно отличался от прогнозов основной массы, его просили пояснить свою позицию, и часто он изменял свои оценки, приближаясь к средним значениям. Эти средние значения и выдавались заказчику как групповое мнение. Надо сказать, что реальные результаты исследования оказались довольно скромными - хотя дата высадки американцев на Луну была предсказана с точностью до месяца, все остальные прогнозы провалились - холодного термоядерного синтеза и средства от рака в XX в. человечество не дождалось.

Однако сама методика оказалась популярной - за последующие годы она использовалась не менее 40 тыс. раз. Средняя стоимость экспертного исследования по методу Дельфи - 5 тыс. долларов США, но в ряде случаев приходилось расходовать и более крупные суммы - до 130 тыс. долларов.

Несколько в стороне от основного русла экспертных оценок лежит *метод сценариев*, применяемый прежде всего для экспертного прогнозирования. Рассмотрим основные идеи технологии сценарных экспертных прогнозов. Экологическое или социально-экономическое прогнозирование, как и любое прогнозирование вообще, может быть успешным лишь при некоторой стабильности условий. Однако решения органов власти, отдельных лиц, иные события меняют условия, и события развиваются по-иному, чем ранее предполагалось. Вполне очевидно, что после первого тура президентских выборов 1996 г. о дальнейшем развитии событий можно было говорить лишь в терминах сценариев: если во втором туре победит Б.Н. Ельцин, то будет то-то и то-то, если же победит Г.А. Зюганов, то события пойдут так-то и так-то.

Метод сценариев необходим не только в социально-экономической или экологической области. Например, при разработке методологического, программного и информационного обеспечения *анализа риска* химико-технологических проектов необходимо составить детальный каталог сценариев аварий, связанных с утечками токсических химических веществ. Каждый из таких сценариев описывает аварию своего типа, со своим индивидуальным происхождением, развитием, последствиями, возможностями предупреждения.

Таким образом, метод сценариев - это метод декомпозиции задачи прогнозирования, предусматривающий выделение набора отдельных вариантов развития событий (сценариев), в совокупности охватывающих все возможные варианты развития. При этом каждый отдельный сценарий должен допускать возможность достаточно точного прогнозирования, а общее число сценариев должно быть обозримо.

Возможность подобной декомпозиции не очевидна. При применении метода сценариев необходимо осуществить два этапа исследования:

- построение исчерпывающего, но обозримого набора сценариев;

- прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы.

Каждый из этих этапов лишь частично формализуем. Существенная часть рассуждений проводится на качественном уровне, как это принято в общественно-экономических и гуманитарных науках. Одна из причин заключается в том, что стремление к излишней формализации и математизации приводит к *искусственному* внесению определенности там, где ее нет по существу, либо к использованию громоздкого математического аппарата. Так, рассуждения на словесном уровне считаются доказательными в большинстве ситуаций, в то время как попытка уточнить смысл используемых слов с помощью, например, теории нечетких множеств приводит к весьма громоздким математическим моделям.

Набор сценариев должен быть обозрим. Приходится исключать различные маловероятные события - прилет инопланетян, падение астероида, массовые эпидемии ранее неизвестных болезней, и т.д. Само по себе создание набора сценариев - предмет экспертного исследования. Кроме того, эксперты могут оценить вероятности реализации того или иного сценария.

Прогнозирование в рамках каждого конкретного сценария с целью получения ответов на интересующие исследователя вопросы также осуществляется в соответствии с описанной выше методологией прогнозирования. При стабильных условиях могут быть применены статистические методы прогнозирования временных рядов. Однако этому предшествует анализ с помощью экспертов, причем зачастую прогнозирование на словесном уровне является достаточным (для получения интересующих исследователя и ЛПР выводов) и не требующим количественного уточнения.

Как известно, при принятии решений на основе *анализа ситуации* (как говорят, при *ситуационном анализе*), в том числе анализе результатов прогнозных исследований, можно исходить из различных критериев. Так, можно ориентироваться на то, что ситуация сложится наихудшим, или наилучшим, или средним (в каком-либо смысле) образом. Можно попытаться наметить мероприятия, обеспечивающие минимально допустимые полезные результаты при любом варианте развития ситуации, и т.д.

Еще один вариант экспертного оценивания - *мозговой штурм*. Организуется он как собрание экспертов, на выступления которых наложено одно, но очень существенное ограничение - нельзя критиковать предложения других. Можно их развивать, можно высказывать свои идеи, но нельзя критиковать! В ходе заседания эксперты, "заражаясь" друг от друга, высказывают все более экстравагантные соображения. Часа через два записанное на магнитофон или видеокамеру заседание заканчивается, и начинается второй этап мозгового штурма - анализ высказанных идей. Обычно из 100 идей 30 заслуживают дальнейшей проработки, из 5-6 дают возможность сформулировать прикладные проекта, а 2-3 оказываются в итоге приносящими полезный эффект - прибыль, повышение экологической безопасности, оздоровление окружающей природной среды и т.п. При этом интерпретация идей - творческий процесс. Например, при обсуждении возможностей защиты кораблей от торпедной атаки была высказана идея: "Выстроить матросов вдоль борта и дуть на торпеду, чтобы изменить ее курс". После проработки эта идея привела к созданию специальных устройств, создающих волны, сбивающиеся торпеду с курса.

Основные стадии экспертного опроса. Более подробно рассмотрим отдельные этапы экспертного исследования. Как показывает опыт, с точки зрения менеджера - организатора такого исследования целесообразно выделять следующие стадии проведения экспертного опроса.

1) *Принятие решения о необходимости проведения экспертного опроса и формулировка Лицом, Принимающим Решения (ЛПР) его цели.* Таким образом, инициатива должна исходить от руководства, что в дальнейшем обеспечит успешное решение организационных и финансовых проблем. Очевидно, что исходный толчок может быть дан докладной запиской одного из сотрудников или дискуссией на совещании, но реальное начало работы - решение ЛПР.

2) *Подбор и назначение ЛПР основного состава Рабочей группы,* сокращенно РГ (обычно - научного руководителя и секретаря). При этом научный руководитель отвечает за организацию и проведение экспертного исследования в целом, а также за анализ собранных материалов и формулировку заключения экспертной комиссии. Он участвует в формировании коллектива экспертов и выдаче задания каждому эксперту (вместе с ЛПР или его представителем). Он сам - высококвалифицированный эксперт и признаваемый другими экспертами формальный и неформальный руководитель экспертной комиссии. Дело секретаря - ведение документации экспертного опроса, решение организационных задач.

3) *Разработка РГ* (точнее, ее основным составом, прежде всего научным руководителем и секретарем) *и утверждение у ЛПР технического задания на проведение экспертного опроса*. На этой стадии решение о проведении экспертного опроса приобретает четкость во времени, финансовом, кадровом, материальном и организационном обеспечении. В частности, формируется Рабочая Группа, в РГ выделяются различные группы специалистов - аналитическая, эконометрическая (специалисты по методам), компьютерная, по работе с экспертами (например, интервьюеров), организационная. Очень важно для успеха, чтобы все эти позиции были утверждены ЛПР.

4) *Разработка аналитической группой РГ подробного сценария (т.е. регламента) проведения сбора и анализа экспертных мнений (оценок).* Сценарий включает в себя прежде всего конкретный вид информации, которая будет получена от экспертов (например, слова, условные градации, числа, ранжировки, разбиения или иные виды объектов нечисловой природы). Например, довольно часто экспертов просят высказаться в свободной форме, ответив при этом на некоторые количество заранее сформулированных вопросов. Кроме того, их просят заполнить формальную карту, в каждом пункте выбрав одну из нескольких градаций. Сценарий должен содержать и конкретные методы анализа собранной информации. Например, вычисление медианы Кемени, статистический анализ люсианов, применение иных методов статистики объектов нечисловой природы и других разделов прикладной статистики (о некоторых из названных методов речь пойдет ниже). Эта работа ложится на эконометрическую и компьютерную группу РГ. Традиционная ошибка - сначала собрать информацию, а потом думать, что с ней делать. В результате, как показывает печальный опыт, информация используется не более чем на 1-2%.

5) *Подбор экспертов* в соответствии с их компетентностью. На этой стадии РГ составляет список возможных экспертов и оценивает степень их пригодности для планируемого исследования..

6) *Формирование экспертной комиссии*. На этой стадии РГ проводит переговоры с экспертами, получает их согласие на работу в экспертной комиссии (сокращенно ЭК). Возможно, часть намеченных РГ экспертов не может войти в экспертную комиссию (болезнь, отпуск, командировка и др.) или отказывается по тем или иным причинам (занятость, условия контракта и др.). ЛПР утверждает состав экспертной комиссии, возможно, вычеркнув или добавив часть экспертов к предложениям РГ. Проводится заключение договоров с экспертами об условиях их работы и ее оплаты.

7) *Проведение сбора экспертной информации.* Часто перед этим проводится набор и обучение интервьюеров - одной из групп, входящих в РГ.

8) Компьютерный *анализ экспертной информации* с помощью включенных в сценарий методов. Ему обычно предшествует введение информации в компьютеры.

9) При применении согласно сценарию экспертной процедуры из нескольких туров - *повторение* двух предыдущих этапов.

10) *Итоговый анализ экспертных мнений, интерпретация полученных результатов* аналитической группой РГ и *подготовка заключительного документа* ЭК для ЛПР.

11) Официальное *окончание* деятельности РГ, в том числе *утверждение ЛПР заключительного документа ЭК*, подготовка и утверждение научного и финансового отчетов РГ о проведении экспертного исследования, оплата труда экспертов и сотрудников РГ, официальное прекращение деятельности (роспуск) ЭК и РГ.

Разберем подробнее отдельные стадии экспертного исследования. Начнем с подбора экспертов: кадры решают все! Каковы эксперты - таково и качество заключения экспертной комиссии.

Подбор экспертов. Проблема подбора экспертов является одной из наиболее сложных в теории и практике экспертных исследований. Очевидно, в качестве экспертов необходимо использовать тех людей, чьи суждения наиболее помогут принятию адекватного решения. Но как выделить, найти, подобрать таких людей? Надо прямо сказать, что *нет методов подбора экспертов, наверняка обеспечивающих успех экспертизы.* Сейчас мы не будем обсуждать проблему существования различных "партий" среди экспертов и обратим внимание на различные иные стороны процедур подбора экспертов.

В проблеме подбора экспертов можно выделить две составляющие - *составление списка возможных экспертов и выбор из них экспертной комиссии в соответствии с компетентностью кандидатов.*

Составление списка возможных экспертов облегчается тогда, когда рассматриваемый вид экспертизы проводится многократно. В таких ситуациях обычно ведется *реестр* возможных экспертов, например, в области государственной экологической экспертизы или судейства фигурного катания, из которого можно выбирать по различным критериям или с помощью датчика (или таблицы) псевдослучайных чисел.

Как быть, если экспертиза проводится впервые, устоявшиеся списки возможных экспертов отсутствуют? Однако и в этом случае у каждого конкретного специалиста есть некоторое представление о том, что требуется от эксперта в подобной ситуации. Для формирования списка есть полезный метод *"снежного кома",* при котором от каждого специалиста, привлекаемого в качестве эксперта, получают определенное количество (обычно 5 - 10) фамилий тех, кто может быть экспертом по рассматриваемой тематике. Очевидно, некоторые из этих фамилий встречались ранее в деятельности РГ, а некоторые - новые. Каждого вновь появившегося опрашивают по той же схеме. Процесс расширения списка останавливается, когда новые фамилии практически перестают встречаться. В результате получается достаточно обширный список возможных экспертов. Метод *"снежного кома"* имеет и недостатки. Число туров до остановки процесса наращивания кома нельзя заранее предсказать. Кроме того, ясно, что если на первом этапе все эксперты были из одного "клана", придерживались в чем-то близких взглядов или занимались сходной деятельностью, то и метод "снежного кома" даст, скорее всего, лиц из этого же "клана". Мнения и аргументы других "кланов" будут упущены. (Здесь речь идет о том, что сообщество специалистов реально разбито на группы, названные выше "кланами", и общение идет в основном внутри "кланов". Неформальная структура науки, к которой относятся "кланы", достаточно сложна для изучения. Отметим здесь, что "кланы" обычно образуются на основе крупных формальных центров (вузов, научных институтов), научных школ.)

Вопрос об оценке компетентности экспертов не менее сложен. Ясно, что успешность участия в предыдущих экспертизах - хороший критерий для деятельности дегустатора, врача, судьи в спортивных соревнованиях, т.е. таких экспертов, которые участвуют в длинных сериях однотипных экспертиз. Однако, увы, наиболее интересны и важны уникальные экспертизы больших проектов, не имеющих аналогов. Использование формальных показателей экспертов (должность, ученые степень и звание, стаж, число публикаций...), очевидно, в современных быстро меняющихся условиях может носить лишь вспомогательный характер, хотя подобные показатели проще всего применять.

Часто предлагают использовать методы самооценки и взаимооценки компетентности экспертов. Обсудим их, начав с метода самооценки, при котором эксперт сам дает информацию о том, в каких областях он компетентен, а в каких - нет. С одной стороны, кто лучше может знать возможности эксперта, чем он сам? С другой стороны, при самооценке компетентности скорее оценивается степень самоуверенности эксперта, чем его реальная компетентность. Тем более, что само понятие *"компетентность"* строго не определено. Можно его уточнять, выделяя составляющие, но при этом усложняется предварительная часть деятельности экспертной комиссии. Достаточно часто эксперт преувеличивает свою реальную компетентность. Например, большинство людей считают, что они хорошо разбираются в политике, экономике, проблемах образования и воспитания, семьи и медицины. На самом деле экспертов (и даже знающих людей) в этих областях весьма мало. Бывают уклонения и в другую сторону, излишне критичное отношение к своим возможностям.

При использовании метода взаимооценки, помимо возможности проявления личностных и групповых симпатий и антипатий, играет роль малая осведомленность экспертов о возможностях друг друга. В современных условиях достаточно хорошее знакомство с работами и возможностями друг друга может быть лишь у специалистов, много лет (не менее 3-4) работающих совместно, в одной комнате, над одной темой. Именно про такие пары можно сказать, что они "*вместе пуд соли съели*". Однако привлечение таких пар специалистов не очень-то целесообразно, поскольку их взгляды из-за схожести жизненного пути слишком похожи друг на друга.

Если процедура экспертного опроса предполагает непосредственное общение экспертов, необходимо учитывать еще ряд обстоятельств. Большое значение имеют их личностные (социально-психологические) качества. Так, один-единственный "*говорун*" может парализовать деятельность всей комиссии на совместном заседании. К срыву могут привести и неприязненные отношения членов комиссии, и сильно различающийся научный и должностной статус членов комиссии. В подобных случаях важно соблюдение регламента работы, разработанного РГ.

Необходимо подчеркнуть, что подбор экспертов – одна из основных функций Рабочей группы, и никакие методики подбора не снимают с нее ответственности. Другими словами, именно на Рабочей группе лежит ответственность за компетентность экспертов, за их принципиальную способность решить поставленную задачу. Важным является требование к ЛПР об утверждении списка экспертов. При этом ЛПР может как добавить в комиссию отдельных экспертов, так и вычеркнуть некоторых из них - по собственным соображениям, с которыми членам РГ и ЭК знакомиться нет необходимости.

Существует ряд нормативных документов, регулирующих деятельность экспертных комиссий в тех или иных областях. Примером является Закон Российской Федерации "Об экологической экспертизе" от 23 ноября 1995 г., в котором регламентируется процедура экспертизы "намечаемой хозяйственной или иной деятельности" с целью выявления возможного вреда, который может нанести рассматриваемая деятельность окружающей природной среде.

О разработке регламента проведения сбора и анализа экспертных мнений. Существует масса методов получения экспертных оценок. В одних с каждым экспертом работают отдельно, он даже не знает, кто еще является экспертом, а потому высказывает свое мнение независимо от авторитетов, "кланов" и отдельных коллег. В других экспертов собирают вместе для подготовки материалов для ЛПР, при этом эксперты обсуждают проблему друг с другом, принимают или отвергают аргументы друг друга, учатся друг у друга, и неверные или недостаточно обоснованные мнения отбрасываются. В одних методах число экспертов фиксировано и таково, чтобы статистические методы проверки согласованности мнений и затем (в случае достаточно хорошей согласованности мнений) их усреднения позволяли принимать обоснованные решения с точки зрения эконометрики. В других - число экспертов растет в процессе проведения экспертизы, например, при использовании метода "снежного кома" для формирования команды экспертов.

В настоящее время *не существует* общепринятой научно обоснованной классификации методов экспертных оценок и тем более - однозначных рекомендаций по их применению. *Попытка силой утвердить одну из возможных точек зрения на классификацию методов экспертных оценок может принести лишь вред.*

Однако для рассказа о многообразии экспертных оценок необходима какая-либо рабочая классификация методов. Одну из таких возможных классификаций мы даем ниже, перечисляя основания, по которым мы делим экспертные оценки.

Один из основных вопросов - что именно должна представить экспертная комиссия в результате своей работы - информацию для принятия решения ЛПР или проект самого решения? От ответа на этот методологический вопрос зависит организация работы экспертной комиссии, и он служит первым основанием для разбиения методов.

ЦЕЛЬ - СБОР ИНФОРМАЦИИ ДЛЯ ЛПР. Тогда Рабочая группа должна собрать возможно больше относящейся к делу информации, аргументов "за" и "против" определенных вариантов решений. Полезен следующий метод постепенного увеличения числа экспертов. Сначала первый эксперт приводит свои соображения по рассматриваемому вопросу. Составленный им материал передается второму эксперту, который добавляет свои аргументы. Накопленный материал поступает к следующему - третьему - эксперту... Процедура заканчивается, когда иссякает поток новых соображений.

Отметим, что эксперты в рассматриваемом методе только поставляют информацию, аргументы "за" и "против", но не вырабатывают согласованного проекта решения. Нет никакой необходимости стремиться к тому, чтобы экспертные мнения были согласованы между собой. Более того, наибольшую пользу приносят эксперты с мышлением, отклоняющимся от массового. Именно от них следует ожидать наиболее оригинальных аргументов.

ЦЕЛЬ - ПОДГОТОВКА ПРОЕКТА РЕШЕНИЯ ДЛЯ ЛПР. Математические методы в экспертных оценках применяются обычно именно для решения задач, связанных с подготовкой проекта решения. При этом зачастую некритически принимают догмы согласованности и одномерности. Эти догмы "кочуют" из одной публикации в другую, поэтому целесообразно их обсудить.

ДОГМА СОГЛАСОВАННОСТИ. Часто без всяких оснований считается, что решение может быть принято лишь на основе согласованных мнений экспертов. Поэтому исключают из экспертной группы тех, чье мнение отличается от мнения большинства. При этом отсеиваются как неквалифицированные лица, попавшие в состав экспертной комиссии по недоразумению или по соображениям, не имеющим отношения к их профессиональному уровню, так и наиболее оригинальные мыслители, глубже проникшие в проблему, чем большинство. Следовало бы выяснить их аргументы, предоставить им возможность для обоснования их точек зрения. Вместо этого их мнением пренебрегают.

Бывает и так, что эксперты делятся на две или более групп, имеющих единые *групповые* точки зрения. Так, известен пример деления специалистов при оценке результатов научно-исследовательских работ на две группы: "теоретиков", явно предпочитающих НИР, в которых получены теоретические результаты, и "практиков", выбирающих те НИР, которые позволяют получать непосредственные прикладные результаты (речь идет о конкурсе НИР в академическом Институте проблем управления (автоматики и телемеханики)).

Иногда заявляют, что в случае обнаружения двух или нескольких групп экспертов (вместо одной согласованной во мнениях) опрос не достиг цели. Это не так! *Цель достигнута - установлено, что единого мнения нет.* Это весьма важно. И ЛПР при принятии решений должен это учитывать. Стремление обеспечить согласованность мнений экспертов любой целой может приводить к сознательному одностороннему подбору экспертов, игнорированию всех точек зрения, кроме одной, наиболее полюбившейся Рабочей группе (или даже "подсказанной" ЛПР).

Часто не учитывают еще одного чисто эконометрического обстоятельства. Поскольку число экспертов обычно не превышает 20-30, то формальная статистическая согласованность мнений экспертов (установленная с помощью тех или иных критериев проверки статистических гипотез) может сочетаться с реально имеющимся разделением экспертов на группы, что делает дальнейшие расчеты не имеющими отношения к действительности. Для примера обратимся к конкретным методам расчетов с помощью коэффициентов конкордации (т.е. - в переводе - согласия) на основе коэффициентов ранговой корреляции Кендалла или Спирмена. Необходимо напомнить, что согласно эконометрической теории положительный результат проверки согласованности таким способом означает ни больше, ни меньше, как отклонение гипотезы о независимости и равномерной распределенности мнений экспертов на множестве всех ранжировок. Таким образом, проверяется нулевая гипотеза, согласно которой ранжировки, описывающие мнения экспертов, являются независимыми случайными бинарными отношениями, равномерно распределенными на множестве всех ранжировок. Отклонение этой нулевой гипотезы по дурной традиции толкуется как согласованность ответов экспертов. Другими словами, мы падаем жертвой заблуждений, вытекающих из своеобразного толкования слов: проверка согласованности в указанном математико-статистическом смысле вовсе не является проверкой согласованности в смысле практики экспертных оценок. (Именно ущербность рассматриваемых математико-статистических методов анализа ранжировок привела группу специалистов к разработке нового эконометрического аппарата для проверки согласованности - непараметрических методов, основанных на т.н. *люсианах* и входящих в современный раздел эконометрики - *статистику нечисловых данных*).

Группы экспертов с близкими методами можно выделить эконометрическими методами кластер-анализа.

МНЕНИЯ ДИССИДЕНТОВ. С целью искусственно добиться согласованности стараются уменьшить влияние мнений экспертов-*диссидентов*, т.е. инакомыслящих по сравнению с большинством. *Жесткий* способ борьбы с диссидентами состоит в игнорировании их мнений, т.е. фактически в их исключении из состава экспертной комиссии. Отбраковка экспертов, как и отбраковка резко выделяющихся результатов наблюдений (выбросов), приводит к процедурам, имеющим плохие или неизвестные статистические свойства. Так, известна *крайняя неустойчивость* классических методов отбраковки выбросов по отношению к отклонениям от предпосылок модели (см., например, учебное пособие [1]).

Мягкий способ борьбы с диссидентами состоит в применении *робастных (устойчивых) статистических процедур*. Простейший пример: если ответ эксперта - действительное число, то резко выделяющееся мнение диссидента сильно влияет на среднее арифметическое ответов экспертов и не влияет на их медиану. Поэтому разумно в качестве согласованного мнения рассматривать медиану. Однако при этом игнорируются (не достигают ЛПР) аргументы диссидентов.

В любом из двух способов борьбы с диссидентами ЛПР лишается информации, идущей от диссидентов, а потому может принять необоснованное решение, которое впоследствии приведет к отрицательным последствиям. С другой стороны, представление ЛПР всего набора мнений снимает часть ответственности и труда по подготовке окончательного решения с комиссии экспертов и рабочей группы по проведению экспертного опроса и перекладывает эти ответственность и труд на плечи ЛПР.

ДОГМА ОДНОМЕРНОСТИ. В устаревшей, а иногда и в современной научно-технической литературе распространен довольно спорный подход так называемой "квалиметрии", согласно которому объект экспертизы всегда можно оценить *одним числом*. Странная идея! *Оценивать человека одним числом приходило в голову лишь на невольничьих рынках*. Вряд ли даже самые рьяные квалиметристы рассматривают книгу или картину как эквивалент числа - ее "рыночной стоимости". Практически все реальные объекты достаточно сложны, а потому сколько-нибудь точно описать их можно лишь с помощью многих и многих чисел, а также математических объектов нечисловой природы.

Вместе с тем нельзя полностью отрицать саму идею поиска обобщенных показателей качества, технического уровня и аналогичных. Так, каждый объект можно оценивать по многим показателям качества. Например, легковой автомобиль можно оценивать по таким показателям:

расход бензина на 100 км пути (в среднем);

надежность (в том числе средняя стоимость ремонта за год);

экологическая безопасность, оцениваемая по содержанию вредных веществ в выхлопных газах;

маневренность (в том числе радиус поворота);

быстрота набора скорости 100 км/час после начала движения; максимальная достигаемая скорость;

длительность сохранения в салоне положительной температуры при низкой наружной температуре (например, минус пятьдесят градусов по Цельсию) и выключенном двигателе;

дизайн (привлекательность и "модность" внешнего вида и отделки салона);

вес, и т.д.

Можно ли свести оценки по этим показателям вместе? Ясно, что определяющей является конкретная ситуация, для которой выбирается автомашина. Максимально достигаемая скорость важна для гонщика, но, как нам представляется, не имеет большого практического значения для водителя рядовой частной машины, особенно в городе с суровым ограничением на максимальную скорость. Для такого водителя важнее расход бензина, маневренность и надежность. Для машин различных служб государственного управления, видимо, надежность важнее, чем для частника, а расход бензина - наоборот. Для районов Крайнего Севера важна теплоизоляция салона, а для южных районов - нет. И т.д.

Таким образом, важна конкретная (узкая) постановка задачи перед экспертами. Но такой постановки зачастую нет. А тогда "игры" по разработке обобщенного показателя качества - например, в виде линейной функции от перечисленных переменных - не могут дать объективных выводов. Альтернативой единственному обобщенному показателю является математический аппарат типа *многокритериальной оптимизации* - множества Парето и т.д.

В некоторых случаях все-таки можно глобально сравнить объекты - например, с помощью тех же экспертов получить упорядочение рассматриваемых объектов - изделий или проектов. Тогда можно ПОДОБРАТЬ коэффициенты при отдельных показателях так, чтобы *упорядочение с помощью линейной функции возможно точнее соответствовало глобальному упорядочению* (например, найти эти коэффициенты методом наименьших квадратов). Наоборот, в подобных случаях НЕ СЛЕДУЕТ оценивать указанные коэффициенты с помощью экспертов. Эта простая идея до сих пор не стала очевидной для отдельных составителей методик по проведению экспертных опросов и анализу их результатов. Они упорно стараются заставить экспертов делать то, что они выполнить *не в состоянии* - указывать веса, с которыми отдельные показатели качества должны входить в итоговый обобщенный показатель.

Эксперты обычно могут сравнить объекты или проекты в целом, но не могут вычленить вклад отдельных факторов. *Раз организаторы опроса спрашивают, эксперты отвечают,* но эти ответы не несут в себе надежной информации о реальности...

ВТОРОЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР - ЧИСЛО ТУРОВ. Экспертизы могут включать один тур, некоторое фиксированное число туров (два, три,…) или неопределенное число туров. Чем больше туров, тем более тщательным является анализ ситуации, поскольку эксперты при этом обычно много раз возвращаются к рассмотрению предмета экспертизы. Но одновременно увеличивается общее время на экспертизу и возрастает ее стоимость. Можно уменьшить расходы, вводя в экспертизу не всех экспертов сразу, а постепенно. Так, например, если цель состоит в сборе аргументов "за" и "против", то первоначальный перечень аргументов может быть составлен одним экспертом. Второй добавит к нему свои аргументы. Суммарный материал поступит к первому и третьему, которые внесут свои аргументы и контраргументы. И так далее - добавляется по одному эксперту на каждый новый тур.

Наибольшие сложности вызывают процедуры с заранее неопределенным числом туров, например, "снежный ком". Часто задают максимально возможное число туров, и тогда неопределенность сводится к тому, придется ли проводить это максимальное число туров или удастся ограничиться меньшим числом.

ТРЕТЬЕ ОСНОВАНИЕ КЛАССИФИКАЦИИ ЭКСПЕРТНЫХ ПРОЦЕДУР - ОРГАНИЗАЦИЯ ОБЩЕНИЯ ЭКСПЕРТОВ. Рассмотрим достоинства и недостатки каждого из элементов шкалы: отсутствие общения - заочное анонимное общение -

заочное общение без анонимности - очное общение с ограничениями - очное общение без ограничений. *При отсутствии общения* эксперт высказывает свое мнение, ничего не зная о других экспертах и об их мнениях. Он полностью независим, что и хорошо, и плохо. Обычно такая ситуация соответствует однотуровой экспертизе. *Заочное анонимное общение*, например, как в методе Дельфи, означает, что эксперт знакомится с мнениями и аргументами других экспертов, но не знает, кто именно высказал то или иное положение. Следовательно, в экспертизе должно быть предусмотрено хотя бы два тура. *Заочное общение без анонимности* соответствует, например, общению по Интернету. Все варианты заочной экспертизы хороши тем, что нет необходимости собирать экспертов вместе, следовательно, находить для этого удобное время и место.

При очных экспертизах эксперты говорят, а не пишут, как при заочных, и потому успевают за то же время сказать существенно больше. *Очная экспертиза с ограничениями* весьма распространена. Это - собрание, идущее по фиксированному регламенту. Примером является военный совет в императорской русской армии, когда эксперты (офицеры и генералы) высказывались в порядке от младшего (по чину и должности) к старшему. Наконец, *очная экспертиза без ограничений* - это свободная дискуссия. Все очные экспертизы имеют недостатки, связанные с возможностями отрицательного влияния на их проведение социально-психологических свойств и клановых (партийных) пристрастий участников, а также неравенства их профессионального, должностного, научного статусов. Представьте себе, что соберутся вместе 5 лейтенантов и 3 генерала. Независимо от того, какая информация имеется у того или иного участника встречи, ход ее предсказать нетрудно: генералы будут беседовать, а лейтенанты - помалкивать. При этом вполне очевидно, что лейтенанты получили образование позже генералов, а потому обладают полезной информацией, которой нет у генералов.

КОМБИНАЦИЯ РАЗЛИЧНЫХ ВИДОВ ЭКСПЕРТИЗЫ. Реальные экспертизы часто представляют собой комбинации различных описанных выше типов экспертиз. В качестве примера рассмотрим защиту студентом дипломного проекта. Сначала идет многотуровая очная экспертиза, проводимая научным руководителем и консультантами, в результате студент подготавливает проект к защите. Затем два эксперта работают заочно - это автор отзыва сторонней организации и заведующий кафедрой, допускающий работу к защите. Обратите внимание на различие задач этих экспертов и объемов выполняемой ими работы - один пишет подробный отзыв, второй росписью на титульном листе проекта разрешает его защиту. Наконец - очная экспертиза без ограничений (для членов ГАК - государственной аттестационной комиссии). Дипломный проект оценивается коллегиально, по большинству голосов, при этом один из экспертов (научный руководитель) знает работу подробно, а остальные - в основном лишь по докладу. Отметим, что мнения экспертов учитываются с весами, а именно, мнения членов ГАК - с весом 1, мнения всех остальных - с весом 0 (совещательный голос). Таким образом, имеем сочетание многотуровой и однотуровой, заочных и очных экспертиз. Подобные сочетания характерны для многих реально проводящихся экспертиз.

1.7 Лекция № 7
Тема: «Общие подходы к исследованию операций на основе программирования»

1.7.2 Вопросы лекции:

1. Математическая постановка задачи линейного программирования.

2. Многокритериальная оптимизация (принцип Парето, лексикографическая оптимизация), вариационные методы получения детерминированных оценок, статистические методы получения оценок, структура и методы принятия решений с использованием различных оценок.

3. Связь между сетевыми задачами и задачами линейного программирования.

1.7.2. Краткое содержание вопросов:

1. Математическая особенность матрицы ограничений транспортной задачи

Под названием транспортная задача объединяется широкий круг задач с единой матетической моделью. Данные задачи относятся к задачам линейного программирования и могут быть решены известным симплексным методом. Однако, обычная транспортная задача имеет большое число переменных и решение ее симплексным методом громозко. С другой стороны матрица системы ограничений транспортной задачи весьма своеобразна, поэтому для ее решения разработаны специальные методы. Эти методы, как и симплексный метод, позволяют найти начальное опорное решение, а затем, улучшая его, получить последовательность опорных решений, которая завершается оптимальным решением.

В практической деятельности часто встречаются задачи, заключающиеся в поиске лучшего (оптимального) решения при наличии различных несводимых друг к другу критериев оптимальности. Например, принятие решения о строительстве дороги в объезд города должно учитывать такие факторы, как выигрыш города в целом по соображениям экологии, проигрыш отдельных предприятий и фирм, например, из-за уменьшения проезжающих через город потенциальных покупателей и многие другие. Если такого рода задачи решаются методами математического программирования, то говорят о задачах ***многокритериальной оптимизации.*** Эти задачи могут носить как линейный, так и нелинейный характер. Поскольку методы решения таких задач излагаются ниже на примере линейных многокритериальных оптимизационных задач, это объясняет рассмотрение этой темы в данной главе учебного пособия.

Задачи многокритериальной оптимизации возникают в тех случаях, когда имеется несколько целей, которые не могут быть отражены одним критерием (например, стоимость и надежность). Требуется найти точку области допустимых решений, которая минимизирует или максимизирует все такие критерии. Если в подобного рода задачах речь идет не о разнородных критериях некоторой системы, а о сопоставлении однородных критериев разных ее подсистем (например, отрасли, группы населения и т.п.), то эти ***задачи называются задачами векторной оптимизации.***

Обозначим 1-й частный критерий через $Z_i(\bar{X})$, где $\bar{X}$- допустимое решение, а область допустимых решений - через ***Q.*** Если учесть, что изменением знака функции всегда можно свести задачу минимизации к задаче максимизации, то кратко задачу многокритериальной оптимизации можно сформулировать следующим образом:

$Z(\bar{X}) = \langle Z_1(\bar{X}), Z_2(\bar{X}), \ldots, Z_m(\bar{X}) \rangle \to \max$(3.28)

$\bar{X} \in Q$.(3.29)

Некоторые частные критерии могут противоречить друг другу, другие действуют в одном направлении, третьи - индифферентны, безразличны друг к другу. Поэтому процесс решения многокритериальных задач неизбежно связан с экспертными оценками как самих критериев, так и взаимоотношений между ними. Известен ряд методов решения задач многокритериальной оптимизации:

- оптимизация одного признанного наиболее важным критерия, остальные критерии при этом играют роль дополнительных ограничений;

- упорядочение заданного множества критериев и последовательная оптимизация по каждому из них (этот подход рассмотрен ниже на примере ***метода последовательных уступок;***

- сведение многих критериев к одному введением экспертных весовых коэффициентов для каждого из критериев таким образом, что более важный критерий получает более высокий вес.

Возвращаясь к задаче многокритериальной оптимизации в общей постановке (3.28), (3.29), отмстим, что в идеальном случае можно вести поиск такого решения, которое принадлежит пересечению множеств оптимальных решений всех однокритериальных задач. Однако такое пересечение обычно оказывается пустым множеством, поэтому приходится рассматривать так называемое переговорное множество ***эффективных решений (оптимальных по Парето).*** Критерий оптимальности итальянского экономиста В. Парето применяется при решении таких задач, когда оптимизация означает улучшение одних показателей при условии, чтобы другие не ухудшались.

Определение 3.1. Вектор $\bar{X}^* \in Q$ называется эффективным (оптимальным по Парето) решением задачи (3.28), (3.29), если не существует такого вектора $\bar{X} \in Q$, что

$Z_i(\bar{X}) \geq Z_i(\bar{X}^*), \quad i = \overline{1,m},$ (3.30)

причем хотя бы для одного значения ***i*** имеет место строгое неравенство.

Множество допустимых решений, для которых невозможно одновременно улучшить все частные показатели эффективности (т.е. улучшить хотя бы один из них, не ухудшая остальных), принято называть ***областью Парето,*** или ***областью компромиссов,*** а принадлежащие ей решения - ***эффективными,*** или ***оптимальными по Парето.***

В общем случае эффективные решения не эквивалентны друг другу, так что про два оптимальных по Парето решения нельзя сказать, какое из них лучше. Поэтому при решении многокритериальных задач необходимо дополнительное изучение эффективных решений. Для этого можно было бы сформулировать некоторый критерий и оптимизировать его на множестве эффективных решений. Однако при этом возникают значительные трудности в связи с тем, что, как правило, область компромиссов не является выпуклой, и полученная задача в общем случае будет задачей невыпуклого программирования. Обычный подход заключается в стремлении "свернуть" частные критерии в один обобщенный скалярный критерий, оптимизация которого приводит к оптимальному решению задачи в целом. Формулировка подходящего обобщенного критерия в зависимости от конкретных условий как раз и является основным вопросом, который изучается в многокритериальной оптимизации.

В некоторых случаях вместо одного обобщенного критерия и решения одной соответствующей задачи скалярной оптимизации предлагается рассматривать последовательность обобщенных критериев и последовательность задач скалярной оптимизации. К сожалению, многие из описанных в литературе подобных процедур не всегда приводят к эффективным решениям.

Рассмотрим один из таких методов решения многокритериальных задач - ***метод последовательных уступок.***

Метод последовательных уступок решения задач многокритериальной оптимизации применяется в случае, когда частные критерии могут быть упорядочены в порядке убывания их важности. Предположим, что все частные критерии максимизируются и пронумерованы в порядке убывания их важности. Находим максимальное значение Z_1^*, первого по важности критерия в области допустимых решений, путем решения однокритериальной задачи

$$Z_1(\bar{X}) \to \max,$$
$$\bar{X} \in Q.$$

Затем, исходя из практических соображений и принятой точности, назначается величина допустимого отклонения 8, > 0 (экономически оправданной уступки) критерия Z, и находится максимальное значение второго критерия ***Z'2*** при условии, что значение первого критерия не должно отклоняться от своего максимального значения более чем на величину допустимой уступки, т.е. решается задача:

$$Z_2(\bar{X}) \to \max,$$
$$Z_1(\bar{X}) \ge Z_1^* - \delta_1,$$
$$\bar{X} \in Q.$$

Снова назначается величина уступки $\delta_2 > 0$ по второму критерию, которая вместе с первой уступкой используется для нахождения условного максимума третьего частого критерия:

$$Z_3(\bar{X}) \to \max,$$
$$Z_1(\bar{X}) \ge Z_1^* - \delta_1,$$
$$Z_2(\bar{X}) \ge Z_2^* - \delta_2,$$
$$\bar{X} \in Q.$$

Аналогичные процедуры повторяются до тех пор, пока не будет выявлено максимальное значение последнего по важности критерия $\boldsymbol{Z_m}$ при условии, что значение каждого из первых $\boldsymbol{m}$ - 1 частных критериев отличается от соответствующего условного максимума не более чем на величину допустимой уступки по данному критерию. Полученное на последнем этапе решение считается оптимальным. Следует заметить, что этот метод не всегда приводит к эффективному решению.

Пример 3.7. Решение задачи многокритериальной оптимизации методом последовательных уступок.

Решение. Пусть задача трехкритериальной оптимизации имеет вид

$$Z_1 = x_1 \quad 2x_2 \to \max \quad (3.31)$$

$$Z_2 = 2x_1 + x_2 \to \max \quad (3.32)$$

$$Z_1 = x_1 \quad 3x_2 \to \max \quad (3.33)$$

$$\left.\begin{aligned} x_1 + x_2 \le 6 \\ 1 \le x_1 \le 3 \\ 1 \le x_2 \le 4 \end{aligned}\right\} \quad (3.34)$$

Заметим, что так как коэффициенты при одних и тех же переменных в данных частных критериях имеют разные знаки, то в заданной области допустимых решений невозможно одновременно улучшить все частные критерии, т.е. в рассматриваемом случае область компромиссов (область Парето) совпадает с областью допустимых решений (3.34).

Для определенности будем считать, что допустимые уступки по первым двум критериям заданы: $\delta_1 = 3$; $\delta_3 = 5/3$.

Максимизируем функцию Z_3 в области допустимых решений, т.е. решаем одну критериальную задачу (3.31), (3.34). Это несложно сделать рассмотренным в главе 2 графическим методом решения задач линейного программирования (рис. 3.3).

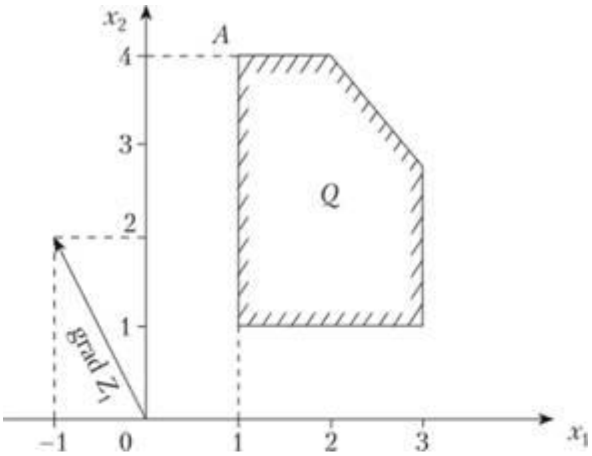


Рис. 3.3

Максимум функции Z_1 при условиях (3.34) достигается в точке ***A*** области ***Q*** с координатами (1; 4), так что в данном случае

$x_1^* = 1;\ \ x_2^*\ \ 4; = \max Z_1\ \ Z_1^*\ \ Z_1(A)\ \ 7$

Переходим к максимизации функции Z, при условиях (3.34) и дополнительном ограничении, позволяющем учесть, что по критерию Z, нельзя уступать более чем на δ_1. Так как в нашем примере $Z_1^* - \delta_1 = 4$, то дополнительное ограничение будет иметь вид

$-x_1 + 2x_2 \geq 4.$(3.35)

Задачу (3.32), (3.34), (3.35) также решаем графически (рис. 3.4).

Получаем, что максимум функции Z_2 при условиях (3.34), (3.35) достигается в точке ***B*** части Q, области Q, так что

$x_1^{**} = 8/3;\ \ x_2^{**}\ \ 10/3;\ \ \max Z_2 = Z_2^*\ \ Z_2(B)\ \ 26/3$

Теперь уступаем по критерию Z_2 на величину уступки 5_2= 5/3 и получаем второе дополнительное ограничение:

$2x_1 + x_2 \geq 7.$(3.36)

Максимизируем функцию Z_3 при условиях (3.34), (3.35) и (3.36). Решение этой задачи представлено на рис 3.5.

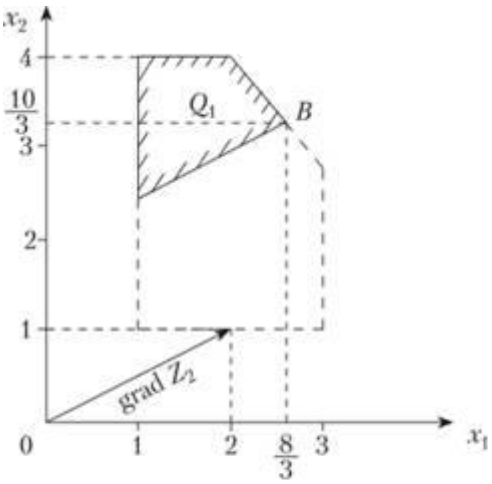


Рис. 3.4

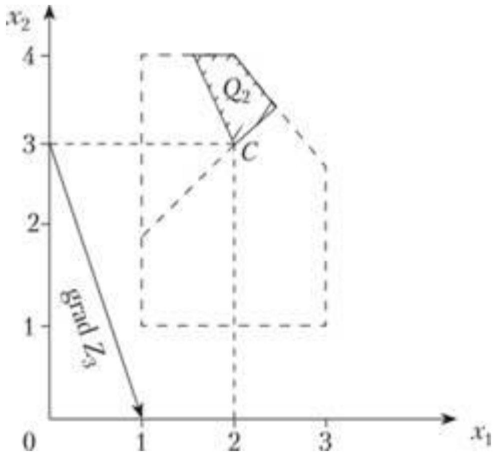


Рис. 3.5

Таким образом, получаем оптимальное решение рассматриваемой трехкритериальной задачи (точка ***C*** на рис. 3.5):

$x_1 = 2$; $\boldsymbol{x_2} = 3$.

Соответствующие значения частных критериев при этом составляют:

$Z_1 = 4$; $Z_2 = 7$; $Z_3 = -7$.

1.8 Лекция № 8
Тема: «Постановка задачи линейного программирования»

1.8.1 Вопросы лекции:

1. Математическая особенность матрицы ограничений транспортной задачи.

2. Метод потенциалов; венгерский метод.

3. Пример решения несбалансированных (сбалансированных), транспортных задач.

1.8.2. Краткое содержание вопросов:

Условие:

Однородный груз сосредоточен у m поставщиков в объемах a_1, a_2, ... a_m.
Данный груз необходимо доставить n потребителям в объемах b_1, b_2 ... b_n.
Известны $\mathbf{C_{ij}}$, i=1,2,...m; j=1,2,...n — стоимости перевозки единиц груза от каждого i-го поставщика каждому j-му потребителю.
Требуется составить такой план перевозок, при котором запасы всех поставщиков вывозятся полностью, запросы всех потребителей удовлетворяются полностью, и суммарные затраты на перевозку всех грузов являются минимальными.

Исходные данные транспортной задачи записываются в виде таблицы:

Таблица 34.1

Потребители

Поставщики

$a_j \backslash b_j$	b_1	b_2	...	b_n
a_1	c_{11}	c_{12}	...	c_{1n}
a_2	c_{21}	c_{22}	...	c_{2n}
...	...	...	...	...
a_m	c_{m1}	c_{m2}	...	c_{mn}

Исходные данные задачи могут быть представлены в виде:

- вектора $A=(a_1,a_2,...,a_m)$ запасов поставщиков
- вектора $B=(b_1,b_2,...,b_n)$ запросов потребителей
- матрицы стоимостей:

$$C = \begin{pmatrix} c_{11} & c_{12} & \dots c_{1n} \\ c_{21} & c_{22} & \dots c_{2n} \\ \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots c_{mn} \end{pmatrix}$$

Математическая модель транспортной задачи

Переменными (неизвестными) транспортной задачи являются x_{ij} , i=1,2,...,m j=1,2,...,n — объемы перевозок от i-го поставщика каждому j-му потребителю.
Эти переменные могут быть записаны в виде матрицы перевозок:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots x_{1n} \\ x_{21} & x_{22} & \dots x_{2n} \\ \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots x_{mn} \end{pmatrix}$$

Так как произведение $C_{ij}*X_{ij}$ определяет затраты на перевозку груза от i-го поставщика j-му потребителю, то суммарные затраты на перевозку всех грузов равны:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}$$

По условию задачи требуется обеспечить минимум суммарных затрат.
Следовательно, целевая функция задачи имеет вид:

$$Z(X) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min$$

Система ограничений задачи состоит из двух групп уравнений.
Первая группа из m уравнений описывает тот факт, что запасы всех m поставщиков вывозятся полностью и имеет вид:

$$\sum_{j=1}^{n} x_{ij} = a_i\ , \ \ i = 1,\ 2,\ \dots,\ m.$$

Вторая группа из n уравнений выражает требование удовлетворить запросы всех n потребителей полностью и имеет вид:

$$\sum_{i=1}^{m} x_{ij} = b_j\ ,\ j = 1, 2, \ldots, n.$$

Учитывая условие неотрицательности объемов перевозок математическая модель выглядит следующим образом:

Математическая модель

$$Z(X) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \to \min, \qquad 1$$

$$\sum_{j=1}^{n} x_{ij} = a_i\ ,\ i = 1, 2, \ldots, m, \qquad 2$$

$$\sum_{i=1}^{m} x_{ij} = b_j\ ,\ j = 1, 2, \ldots, n, \qquad 3$$

$$x_{ij} \geq 0,\ i = 1, 2, \ldots, m; j = 1, 2, \ldots, n. \qquad 4$$

В рассмотренной модели транспортной задачи предполагается, что суммарные запасы поставщиков равны суммарынм запросам потребителей, т.е.:

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

Такая задача называется задачей с **правильным балансом**, а модель задачи **закрытой**. Если же это равенство не выполняется, то задача называется задачей с **неправильным балансом**, а модель задачи — **открытой**.

Математическая формулировка транспортной задачи такова: найти переменные задачи $X=(x_{ij})$, $i=1,2,...,m$; $j=1,2,...,n$, удовлетворяющие системе ограничений (цифра 2 на математической модели) , (3), условиям неотрицательности (4) и обеспечивающие минимум целевой функции (1)

Пример 34.1

Составить математическую модель транспортной задачи, исходные данные которой приведены в таблице 34.2

Таблица 34.2

a_i \ b_j	20	30	40
40	3	5	7
50	4	6	10

Решение:

1. **Вводим переменные задачи (матрицу перевозок):**

$$X = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{pmatrix}.$$

2. **Записываем матрицу стоимостей:**

$$C = \begin{pmatrix} 3 & 5 & 7 \\ 4 & 6 & 10 \end{pmatrix}$$

3. **Целевая функция задачи равняется сумме произведений всех соответствующих элементов матриц C и X.**

$$Z(X) = 3x_{11} + 5x_{12} + 7x_{13} + 4x_{21} + 6x_{22} + 10x_{23}.$$

Данная функция, определяющая суммарные затраты на все перевозки, должна достигать минимального значения.

4. **Составим систему ограничений задачи.**
Сумма всех перевозок, стоящих в первой строке матрицы X, должна равняться запасам первого поставщика, а сумма перевозок во второй строке матрицы X равняться запасам второго поставщика:

$$x_{11} + x_{12} + x_{13} = 40,$$

$$x_{21} + x_{22} + x_{23} = 50.$$

Это означает, что запасы поставщиков вывозятся полностью.

Суммы перевозок, стоящих в каждом столбце матрицы X, должны быть равны запросам соответствующих потребителей:

$$x_{11} + x_{21} = 20,$$

$$x_{12} + x_{22} = 30,$$

$$x_{13} + x_{23} = 40.$$

Это означает, что запросы потребителей удовлетворяются полностью.

Необходимо также учитывать, что перевозки не могут быть отрицательными:

$$x_{ij} \geq 0,\ i = 1, 2, ..., m; j = 1, 2, ..., n.$$

Ответ: Таким образом, математическая модель рассматриваемой задачи записывается следующим образом:
Найти переменные задачи, обеспечивающие минимум целевой функции (1) и удовлетворяющие системе ограничений (2) и условиям неотрицательности (3).

$$Z(X) = 3x_{11} + 5x_{12} + 7x_{13} + 4x_{21} + 6x_{22} + 10x_{23} \quad (1)$$

$$\begin{cases} x_{11} + x_{12} + x_{13} = 40, \\ x_{21} + x_{22} + x_{23} = 50, \\ x_{11} + x_{21} = 20, \\ x_{12} + x_{22} = 30, \\ x_{13} + x_{23} = 40 \end{cases} \quad (2)$$

$$x_{ij} \geq 0,\ i = 1, 2, ..., m; j = 1, 2, ..., n. \quad (3)$$

Метод потенциалов

Метод потенциалов используется для решения транспортной задачи. Основой вычислительного процесса при улучшении опорного плана является определение критерия оптимальности d_{ij}:

$$d_{ij} = c_{ij}^{*} - c_{ij},$$

где c_{ij} - затраты (истинные тарифы), связанные с доставкой одной единицы груза из i-того пункта отправления в j-ый пункт назначения; c_{ij}^{*} - расчётные затраты (косвенные тарифы), связанные с доставкой одной единицы груза из i-того пункта отправления в j-ый пункт назначения, определяемые для тех клеток опорного плана, ресурсы в которые не распределены (для незаполненных клеток).

Алгоритм метода потенциалов

План транспортной задачи является оптимальным, если для всех свободных клеток таблицы перевозок значение критерия оптимальности d_{ij}<=0. Если для всех свободных клеток таблицы перевозок критерий оптимальности d_{ij}<0, то этот план перевозок является единственным. Если для некоторых свободных клеток таблицы перевозок критерий оптимальности d_{ij}=0, то этот оптимальный план перевозок не является единственным. Наконец, если имеются свободные клетки, для которых критерий оптимальности d_{ij}>0, то полученный план перевозок не является оптимальным. Алгоритм метода потенциалов состоит в следующем: каждому поставщику A_i ставится в соответствие некоторое число *u*, которое называется потенциалом A_i-того поставщика; каждому потребителю B_j ставится в соответствие некоторое число *v*, которое называется потенциалом B_j-того потребителя. Для каждой заполненной клетки, т. е. для каждой базисной переменной строится соотношение:

$$u_i+v_j=c_{ij}$$

Получаем систему с числом уравнений, равным количеству базисных переменных. Из этой системы определяем неизвестные потенциалы u_i и v_j, полагая u_i=0. Для каждой незаполненной клетки, т. е. для каждой небазисной переменной, рассчитываются косвенные тарифы c_{ij}^* по формуле: $c_{ij}^* = u_i+v_j$. Затем полученный план проверяют на оптимальность по критерию оптимальности d_{ij}. Если для каждой незаполненной клетки выполняется условие: $d_{ij}<=c_{ij}^*<=0$, то исходный план является оптимальным. Если некоторые d_{ij}>0, то необходимо перейти к новому плану путем перемещения перевозки в клетку, отвечающую условию max(d_{ij}). Если таких клеток несколько, то выбирают любую из них.

Для правильного перемещения перевозок, чтобы не нарушить ограничения задачи, строят так называемый цикл, т. е. замкнутый многоугольник, соединяющий выбранную клетку с ней же самой и проходящий через заполненные клетки.

Цикл строится следующим образом: вычёркиваются (мысленно) все строки и столбцы, содержащие ровно одну заполненную клетку, при этом считается, что выбранная клетка без поставки является заполненной; все оставшиеся после вычеркивания клетки составляют цикл и лежат в его углах, они соединяются ломаной линией.

В каждую клетку цикла, начиная с незаполненной, поочередно вписывают знаки “+” и “-“. В клетках с отрицательными знаками выбирается минимальная величина поставки, обозначаемая как *D*. В те вершины, которые имеют знак “+” прибавляется поставка *D*, а в вершинах со знаком “-“ поставки уменьшаются на величину *D*. При этом суммы поставок по строкам и столбцам не изменяются. В результате клетка, для которой строился цикл, станет занятой, а в одной из бывших занятых клеток окажется нулевая поставка и её надо объявить свободной. Общее количество заполненных клеток не изменится, следовательно, новый план перевозок является невырожденным.

Если в результате пересчета одновременно в нескольких ранее занятых клетках цикла поставки примут нулевые значения, то свободной объявляется лишь одна из них. Остальные считаются условно занятыми с нулевыми поставками.

Значения переменных, включенных в цикл, после пересчета переносятся в новую таблицу без изменений. Полученный новый план проверяется на оптимальность. Такое улучшение плана можно проводить неоднократно до тех пор пока все критерии для незаполненных клеток окажутся $d_{ij}<=0$. Затем вычисляется оптимальная стоимость перевозок.

Венгерский метод

Специфические особенности задач о назначениях послужили поводом к появлению эффективного венгерского метода их решения. Основная идея венгерского метода заключается в переходе от исходной квадратной матрицы стоимости C к эквивалентной ей матрице Cэ с неотрицательными элементами и системой n независимых нулей, из которых никакие два не принадлежат одной и той же строке или одному и тому же столбцу. Для заданного n существует n! допустимых решений. Если в матрице назначения X расположить n единиц так, что в каждой строке и столбце находится только по одной единице, расставленных в соответствии с расположенными n независимыми нулями эквивалентной матрицы стоимости Cэ, то получим допустимые решения задачи о назначениях.

Следует иметь в виду, что для любого недопустимого назначения соответствующая ему стоимость условно полагается равной достаточно большому числу M в задачах на минимум. Если исходная матрица не является квадратной, то следует ввести дополнительно необходимое количество строк или столбцов, а их элементам присвоить значения, определяемые условиями задачи, возможно после редукции, а доминирующие альтернативы дорогие или дешевые исключить.

1.9 Лекция №9 (2 часа)

Тема: «Симплексный метод решения задач линейного программирования»

1.9.1 Вопросы лекции:

1. Основные определения в территории графов и в сетевых задачах (граф ориентированный, граф и ориентированный цикл, дерево, цепь, путь, контур, петля).

1.9.2 Краткое содержание вопросов:

1. Основные определения в территории графов и в сетевых задачах (граф ориентированный, граф и ориентированный цикл, дерево, цепь, путь, контур, петля).

Если рассматривать логистику как задачу нахождения оптимального перемещения, то с этой точки зрения – каждый человек в течении дня на уровне подсознания неоднократно решает такую задачу, перемещаясь:
– из дома на работу и обратно;
– с работы в театр, в кино, в кафе и т.п.;
– в столовую во время обеда, по помещениям и учреждениям во время работы, и т.д.
– даже первобытный человек во время охоты, решал данную задачу в процессе погони за добычей!
В своей общественно-практической деятельности человек также каждодневно решает задачи логистики оптимального перемещения:
– грузов;

– войск;
– готовой продукции и комплектующих в процессе её производства;
– пассажиров и многого другого.

Логистика – как наука известна со времен Римской империи, где: “Логист” – снабженец в армии.

В настоящее время – основная задача логистики, это:

Линейная транспортная задача нахождения способов и путей наиболее оптимальной и быстрой доставки грузов, товаров, пассажиров и т.п. к пунктам назначения.*

К началу 20-го века Логистика, как наука, получила в арсенал своих научных средств новый математический аппарат – теорию графов

* Данное определение логистики не охватывает все её задачи и выбрано как наиболее простое, поскольку к настоящему времени дано более 200 определений, многие из которых противоречивы к другим.

2. История возникновения теории графов

Датой рождения теории графов принято считать 1736 год, когда Леонард Эйлер сформулировал первые теоремы новой теории, решая задачу "О кенигсбергских мостах", и написал об этом в письме итальянскому математику и инженеру Мариони 13 марта 1736 г., чем и закрепил за собой приоритет автора новой теории.
Однако: Термин «граф» впервые ввел в математику венгерский математик Денеш Кениг в 1936 году.

В настоящее время:

– в некоторых отраслях прикладной деятельности вместо понятия "граф" используется термин "сеть": сетевой график; сеть железных дорог и т.п.;
– вместо термина "вершина графа" при использовании "сетевой" терминологии используется термин "сетевой узел" или просто – "узел".

"Задача о кенигсбергских мостах".

В 1736 г. город Калининград назывался Кёнигсбергом. Он был расположен на берегах р. Прегель, и на двух островах между ними. Четыре образовавшихся участка суши (правый и левый берег, и два острова) соединяло семь мостов так (рис.1). В задаче предлагалось составить маршрут для почтальона, полицейского или туриста, чтобы они побывали во всех районах города, и при этом прошли по каждому мосту только один раз – в настоящее время такой путь называется "Эйлеровым путём".

В числе сформулированных Эйлером теорем была теорема о том, что: **Граф с более чем двумя нечётными степенями вершин невозможно начертить или обойти все его вершины так, чтобы пройти по каждому ребру графа только один раз.**

При этом степень вершины он определил равной числу линий – "ребер", связанных с данной вершиной или "сетевым узлом" графа.

На рис. 2 показано изображение графа, для "Задачи о кенигсбергских мостах".

Вершины графа соответствуют определённому району города, а ребра – мостам через реку.
Из рисунка видно, что все четыре вершины графа – имеют нечетные степени: вершины А, В, Г – имеют степень 3, вершина Б – степень 5, т.е. "Эйлерового пути" у данного графа не существует!

На рис. 3 приведены примеры изображений графов, для которых существуют "Эйлеровы пути".

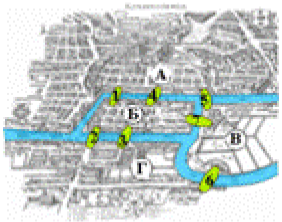


Рис. 1.

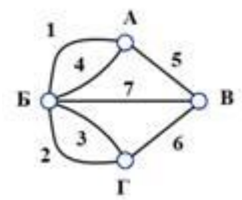


Рис. 2.

Эйлеровы Графы:

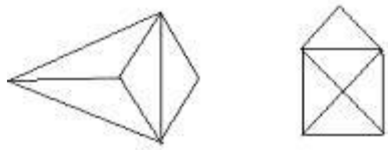


Рис. 3.

Далее, вплоть до начала 20-го века теория графов развивалась в основном в виде формулирования новых теорем, сформированных по результатам решения различных "головоломных задач".

Серьезное развитие теория графов получила в связи с возникновением массового крупносерийного производства, общим всплеском науки и технологий в первой половине 20-го века.

3. Области применения графов.

В настоящее время теория графов нашла очень широкое применение в:

1. В микроэлектронике – при разработке топологии микросхем.
2. При разработке сложных электрических схем и схем их монтажа в электротехнических шкафах, в электрощитах.
3. В химии – при разработке новых сложных молекулярных соединений и т.п.
4. Физике – при описании и анализе схем развития квантовых процессов.
5. При разработке коммуникационных систем различного назначения.
6. В транспортных системах – как для изучения самих систем, так и при составлении оптимальных маршрутов доставки грузов – логистика.
7. В информатике и программирование – при разработке алгоритмов расчетов, программ.
8. В экономике и планировании – в виде сетевых графиков.
и т.д.

4. Основные понятия теории графов.

Графом, или точнее: "**Изображением графа**" – называется конечное множество точек, соединенных, **как правило**, друг с другом любыми линиями. Данные точки называются вершинами графа или узлами сети, а соединяющие их линии – рёбрами. Если, точка не соединена ни с одной другой точкой – она называется "**висячей**". **Степень вершины** графа равна количеству линий или ребер, связанных с данной точкой, вершиной, узлом. p(N)
С точки зрения решения задачи, связанной с конкретным изображением графа: **Граф** непосредственнл – **это решение поставленной задачи** в виде конкретной схемы пути по вершинам и ребрам изображения графа. Каждая задача может иметь одно и более решений, либо не иметь решения.
Эйлеров граф – это такой граф, в котором существует путь по всем его ребрам такой, чтобы каждое ребро было пройдено только один раз.
Гамильтонов путь – путь через все вершины графа наиболее коротким путем. Пути могут быть замкнутыми, т.е. начинаться и заканчиваться в одной точке, либо разомкнутыми.
Ориентированные графы – граф, имеющий ребра по которым можно перемещаться только в одном направлении.
Неориентированные графы – в котором по всем ребрам можно перемещаться в любую сторону.

Элементарные теоремы:

$$\sum_{i=1}^{\text{Все}} p(i) = 2R$$

Теорема 1.

Сумма степеней вершин графа – четное и равна удвоенному числу ребер графа.
Доказательство: Каждое ребро соединяет две вершины (например A и B), и будет дважды учтено при определении степеней узлов, поскольку: p(A) = p(B) =1 и p(A) + p(B) = 2.

Теорема 2.

Число узлов с нечетной степенью любого графа – четное.

Доказательство: Из Теоремы 1 мы знаем, что сумма четных и нечетных степеней вершин – четная: $\sum p_ч + \sum p_н = 2R$, а сумма четных чисел $\sum p_ч$ – всегда четная. Значит и $\sum p_н$ – четная.

Теорема 3.

У Графа с количеством вершин нечетной степени больше 2-х Эйлерова пути нет.

Доказательство: выполнено Эйлером в 1736 г.

5. Наиболее известные и "яркие" задачи теории графов:

1. **Теорема о четырёх красках** утверждает, что всякую расположенную на сфере карту можно раскрасить четырьмя красками так, чтобы любые две области, имеющие общий участок границы, были раскрашены в разные цвета. Теорема была сформулирована в 1852 году, однако доказать ее долгое время не удавалось. В течение этого времени было предпринято множество попыток как доказательства, так и опровержения, и эта задача носила название *проблемы четырёх красок*. Для простых карт достаточно и трёх цветов, а четвёртый цвет начинает требоваться, например, тогда, когда имеется одна область, окруженная нечетным числом других, которые соприкасаются друг с другом, образуя цикл. **Теорема о пяти красках**, утверждающая, что достаточно пяти цветов, имела короткое несложное доказательство и была доказана в конце XIX века, но доказательство теоремы для случая четырёх цветов столкнулось со значительными трудностями и была доказана лишь в 1976 году. Это была первая крупная математическая теорема, доказанная с помощью компьютера. Доказательство этого факта заняло сотни страниц и не все математики признали его. В 1997 году, и в 2005 году были найдены более простые доказательства, также проделанные с помощью компьютера.

2. **Задача коммивояжера** – в которой необходимо посетить каждый город в пределах определенной территории и возвратиться в пункт отправления, причем так, чтобы путь был как можно короче. С современной точки зрения – это типичная задача логистики, была придумана исключительно для развлечения.

3. **Задача о ходе коня** – задача о нахождении маршрута шахматного коня, проходящего через все поля стандартной шахматной доски по одному разу. Как оказалось: Количество всех замкнутых маршрутов коня (гамильтоновых циклов) без учёта направления обхода равно:
13 267 364 410 532.Количество всех незамкнутых маршрутов (с учётом направления обхода) равно: 19 591 828 170 979 904. Наиболее красивое решение (рис. 4) получено при помощи "Шахматного компьютера". Однако с точки зрения логистики наиболее интересным кажется маршрут, найденный К. Я. Янишем, в котором конь сначала обходит одну половину доски, а затем вторую.

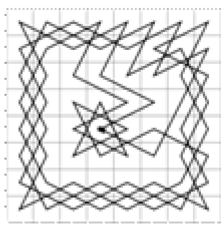


Рис. 4.

6. Применение Теории Графов в логистике.

Как уже указано выше – нахождение замкнутого Гамильтонова пути – это фактически и есть **главная задача современной логистики**.

2. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ВЫПОЛНЕНИЮ ЛАБОРАТОРНЫХ РАБОТ (не предусмотрены рабочей программой)

3. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

3.1 Практическое занятие № 1

Тема: «Сетевые потоковые задачи»

Задача. Выяснить пропускную способность сети

Решение. Опишем сеть матрицей пропускных способностей

$K_0=$

	s	1	2	3	4	5	6	t
s	0	1	4	0	0	4	0	0
1	1	0	0	1	1	0	2	0
2	4	0	0	2	0	0	0	1
3	0	1	2	0	0	0	3	0
4	0	1	0	0	0	3	1	0
5	4	0	0	0	3	0	0	1
6	0	2	0	3	1	0	0	6
t	0	0	1	0	0	1	6	0

Для начала возьмем исходный поток F_0 в котором одна единица перевозится по верхнему пути, одна по нижнему, одна по среднему. Получим матрицу потока F_1.

$F_1=$

	s	1	2	3	4	5	6	t
s		1	1			1		
1	-1						1	
2	-1							1
3								
4								
5	-1							1
6		-1						1
t			-1			-1	-1	

Найдем пути, которые не насыщены потоком F_1. Для этого вычтем F_1 из K_0, то есть $K_1=K_0 - F_1$.

Отыщем все узлы которых можно достичь идя из s по ненасыщенным ребрам. Это соответствует положительным элементам первой строки матрицы K_1.

$K_1=$

	s	1	2	3	4	5	6	t
s	0	0	(3)	0	0	3	0	0
1	2	0	0	1	1	0	1	0
2	[5]	0	0	(2)	0	0	0	0
3	0	1	[2]	0	0	0	(3)	0
4	0	1	0	0	0	3	1	0
5	5	0	0	0	3	0	0	0

	s	1	2	3	4	5	6	t
6	0	3	0	[3]	1	0	0	(5)
t	0	0	2	0	0	2	[7]	0

(s, 2)(2, 3)(3, 6)(6, t) - (*) или (s, 5)(5, 4)(4, 6)(6, t).

Путь (*). Образующие его элементы в K_1 выделены ( ), а элементы соответствующие ребрам противоположного направления выделены [ ].Минимальное значение пропускных способностей ребер ( ) равно 2, поэтому по выделенному пути можно направить поток мощности 2.

K_2 можно получить вычитанием 2 из ( ) и дополнением [ ].

K_2=

	s	1	2	3	4	5	6	t
s		0	1			(3)		
1	2			1	1		1	
2	7			0				0
3		1	4				1	
4		1				[3]	(1)	
5	[5]				(3)			0
6		3		5	[1]			(3)
t			2			2	[9]	

K_2: (s, 5)(5, 4)(4, 6)(6, t), min{3,3,1,3}=1.

K_3: (s, 5)(5, 4)(4, 1)(1, 6)(6, t), min{2,2,1,1,2}=1.

K_3=

	s	1	2	3	4	5	6	t
s		0	1			(2)		
1	2			1	[1]		(1)	
2	7			0				0
3		1	4				1	
4		(1)				[4]	0	
5	[6]				(2)			0
6		[3]		5	2			(2)
t			2			2	[10]	

В выделяем ( ) все узлы, которых можно достичь из s. Они не включают t, поэтому вычисления закончены. Единственными узлами, которых можно достичь из s в ( ) являются узлы (s), (2), (4), (5), следовательно, они образуют множество S минимального сечения.
S={s, 2, 4, 5}, S'={1, 3, 6, t}.
Найдем пропускную способность сечения: C(S, S')=c(s, 1)+c(1, 4)+c(2, 3)+c(2, t)+c(4, 6)+c(5,t)=1+1+2+1+1+1=7.
Матрицу максимального потока F находим по формуле $F=K_0 - K_4$.

K_4=

	s	1	2	3	4	5	6	t
s		0	(1)			(1)		
1	2			1	2		0	
2	7			0				0
3		1	4				1	
4		1				5	0	
5	7				(1)			0
6		4		5	2			1
t			2			2	11	

В результате получили, что общий сток, выходящий из s, равен сумме чисел первой строки и равен общему стоку, который представляет собой сумму элементов столбца t. Мощность потока равна7, что совпадает с пропускной способностью минимального сечения. Это может служить проверкой вычислений.

F=

	s	1	2	3	4	5	6	t
s		1	3			3		
1	-1				-1		2	
2	-3			2				1
3			-2				2	
4		1				-2	1	
5	-3				2			1
6		-2		-2	-1			5
t			-1			-1	-5	

s →1 →6 →4 (1)

s →2 →t (1)
s →2 →3 →6 → t (2)
s →5 →t (1)
s →5 →4 →6 → t (1)
s →5 →4 →1 → 6 → t (1)

3.2 Практическое занятие №2

Тема: «Классификация систем массового

обслуживания»

Системы массового обслуживания - это такие системы, в которые в случайные моменты времени поступают заявки на обслуживание, при этом поступившие заявки обслуживаются с помощью имеющихся в распоряжении системы каналов обслуживания.

С позиции моделирования процесса массового обслуживания ситуации, когда образуются очереди заявок (требований) на обслуживание, возникают следующим образом. Поступив в обслуживающую систему, требование присоединяется к очереди других (ранее поступивших) требований. Канал обслуживания выбирает требование из находящихся в очереди, с тем чтобы приступить к его обслуживанию. После завершения процедуры обслуживания очередного требования канал обслуживания приступает к обслуживанию следующего требования, если таковое имеется в блоке ожидания.

Цикл функционирования системы массового обслуживания подобного рода повторяется многократно в течение всего периода работы обслуживающей системы. При этом предполагается, что переход системы на обслуживание очередного требования после завершения обслуживания предыдущего требования происходит мгновенно, случайные моменты времени.

Примерами систем массового обслуживания могут служить:

1. посты технического обслуживания автомобилей;

2. посты ремонта автомобилей;

3. персональные компьютеры, обслуживающие поступающие заявки или требования на решение тех или иных задач;

4. станции технического обслуживания автомобилей;

5. аудиторские фирмы;

6. отделы налоговых инспекций, занимающиеся приемкой и проверкой текущей отчетности предприятий;

7. телефонные станции и т. д.

Основными компонентами системы массового обслуживания любого вида являются:

- входной поток поступающих требований или заявок на обслуживание;

- дисциплина очереди;

- механизм обслуживания.

Входной поток требований. Для описания входного потока требуется задать вероятностный закон, определяющий последовательность моментов поступления требований на обслуживание и указать количество таких требований в каждом очередном поступлении. При этом, как правило, оперируют понятием «вероятностное распределение моментов поступления требований». Здесь могут поступать как единичные, так и групповые требования (требования поступают группами в систему). В последнем случае обычно речь идет о системе обслуживания с параллельно-групповым обслуживанием.

Дисциплина очереди - это важный компонент системы массового обслуживания, он определяет принцип, в соответствии с которым поступающие на вход обслуживающей системы требования подключаются из очереди к процедуре обслуживания.

Механизм обслуживания определяется характеристиками самой процедуры обслуживания и структурой обслуживающей системы. К характеристикам процедуры обслуживания относятся: продолжительность процедуры обслуживания и количество требований, удовлетворяемых в результате выполнения каждой такой процедуры. Для аналитического описания характеристик процедуры обслуживания оперируют понятием «вероятностное распределение времени обслуживания требований».

Следует отметить, что время обслуживания заявки зависит от характера самой заявки или требований клиента и от состояния и возможностей обслуживающей системы. В ряде случаев приходится также учитывать вероятность выхода обслуживающего прибора по истечении некоторого ограниченного интервала времени.

Структура обслуживающей системы определяется количеством и взаимным расположением каналов обслуживания (механизмов, приборов и т. п.). Прежде всего следует подчеркнуть, что система обслуживания может иметь не один канал обслуживания, а несколько; система такого рода способна обслуживать одновременно несколько требований. В этом случае все каналы обслуживания предлагают одни и те же услуги, и, следовательно, можно утверждать, что имеет место параллельное обслуживание.

Разнотипные каналы массового обслуживания

Система обслуживания может состоять из нескольких разнотипных каналов обслуживания, через которые должно пройти каждое обслуживаемое требование, т. е. в обслуживающей системе процедуры обслуживания требований реализуются последовательно. Механизм обслуживания определяет характеристики выходящего (обслуженного) потока требований.

Простейшей одноканальной моделью с вероятностными входным потоком и процедурой обслуживания является модель, характеризуемая показательным распределением как длительностей интервалов между поступлениями требований, так и длительностей обслуживания. При этом плотность распределения длительностей интервалов между поступлениями требований имеет вид

$$f_1(x) = \lambda \cdot e^{-\lambda t},$$

где λ - интенсивность поступления заявок в систему

Плотность распределения длительностей обслуживания:

$$f_2(x) = \mu \cdot e^{-\mu t},$$

где μ - интенсивность обслуживания

Потоки заявок и обслуживании простейшие.

Пусть система работает с отказами. Необходимо определить абсолютную и относительную пропускную способность системы.

Представим данную систему массового обслуживания в виде графа (рис. 1), у которого имеются два состояния:

$S0$ - канал свободен (ожидание);

$S1$ - канал занят (идет обслуживание заявки).

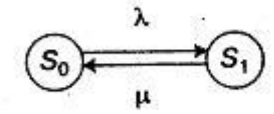


Рис. 1. Граф состояний одноканальной СМО с отказами

Обозначим вероятности состояний: $P0(t)$ - вероятность состояния «канал свободен»; $P1(t)$ - вероятность состояния «канал занят». По размеченному графу состояний (рис. 1) составим систему дифференциальных уравнений Колмогорова для вероятностей состояний:

$$\begin{cases} \dfrac{dP_0(t)}{dt} = -\lambda \cdot P_0(t) + \mu \cdot P_1(t) \\ \dfrac{dP_1(t)}{dt} = -\mu \cdot P_1(t) + \lambda \cdot P_0(t). \end{cases}$$

Система линейных дифференциальных уравнений (4.3) имеет решение с учетом нормировочного условия $P0(t) + P1(t) = 1$. Решение данной системы называется неустановившимся, поскольку оно непосредственно зависит от t и выглядит следующим образом:

$$P_0(t) = \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t} + \frac{\mu}{\lambda + \mu},$$

$P1(t) = 1 - P0(t) = 1$.

Нетрудно убедиться, что для одноканальной СМО с отказами вероятность $P0(t)$ есть не что иное, как относительная пропускная способность системы q.

Действительно, $P0$ - вероятность того, что в момент t канал свободен и заявка, пришедшая к моменту t, будет обслужена, а следовательно, для данного момента времени t среднее отношение числа обслуженных заявок к числу поступивших также равно $P0(t)$, т. е.

$q = P0(t)$,

По истечении большого интервала времени (при $t \to \infty$) достигается стационарный (установившийся) режим:

$$q = P_0 = \frac{\mu}{\mu + \lambda},$$

Зная относительную пропускную способность, легко найти абсолютную. Абсолютная пропускная способность (А) - среднее число заявок, которое может обслужить система массового обслуживания в единицу времени:

$$A = \lambda \cdot q = \frac{\lambda\mu}{\lambda + \mu}.$$

Вероятность отказа в обслуживании заявки будет равна вероятности состояния «канал занят»:

$$P_{отк} = P_1 = 1 - P_0 = 1 - \frac{\mu}{\lambda + \mu} = \frac{\lambda}{\lambda + \mu}.$$

Данная величина *Ротк* может быть интерпретирована как средняя доля не обслуженных заявок среди поданных.

Система массового обслуживания имеет один канал. Входящий поток заявок на обслуживание - простейший поток с интенсивностью λ. Интенсивность потока обслуживания равна μ (т. е. в среднем непрерывно занятый канал будет выдавать μ обслуженных заявок). Длительность обслуживания - случайная величина, подчиненная показательному закону распределения. Поток обслуживании является простейшим пуассоновским потоком событий. Заявка, поступившая в момент, когда канал занят, становится в очередь и ожидает обслуживания.

Предположим, что независимо от того, сколько требований поступает на вход обслуживающей системы, данная система (очередь + обслуживаемые клиенты) не может вместить более N-требований (заявок), т. е. клиенты, не попавшие в ожидание, вынуждены обслуживаться в другом месте. Наконец, источник, порождающий заявки на обслуживание, имеет неограниченную (бесконечно большую) емкость.

Граф состояний СМО в этом случае имеет вид, показанный на рис. 2.

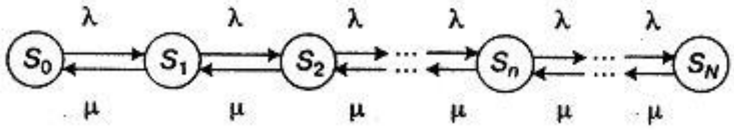


Рис. 2. Граф состояний одноканальной СМО с ожиданием (схема гибели и размножения)

Состояния СМО имеют следующую интерпретацию:

*S*0 - «канал свободен»; *S*1 - «канал занят» (очереди нет); *S*2 - «канал занят» (одна заявка стоит в очереди); *Sn* - «канал занят» (n -1 заявок стоит в очереди); *SN* - «канал занят» (N - 1 заявок стоит в очереди).

Стационарный процесс в данной системе будет описываться следующей системой алгебраических уравнений:

$$\begin{cases} -\psi \cdot P_0 + P_1 = 0, & n = 0 \\ \dots\dots\dots\dots\dots\dots & \\ -(1-\psi) \cdot P_n + P_{n+1} + \psi \cdot P_{n-1} = 0, & 0 < n < N \\ \dots\dots\dots\dots\dots\dots & \\ -P_N + \psi \cdot P_{N-1} = 0, & n = N \end{cases},$$

где $\psi = \dfrac{\lambda}{\mu}$; *n* – номер состояния.

Решение приведенной выше системы уравнений (4.10) для нашей модели СМО имеет вид

$$P_n = \begin{cases} \left(\dfrac{1-\psi}{1-\psi^{N+1}}\right)\cdot \psi^n, & \psi \neq 1, \quad n = 0,1,2,\ldots,N \\ \dfrac{1}{(N+1)}, & \psi = 1. \end{cases}$$

$$P_0 = \frac{1-\psi}{1-\psi^{N+1}}$$

Тогда

$$P_n = \begin{cases} P_0 \cdot \psi^n, & \psi \neq 1, \quad n = 0,1,2,\ldots,N \\ \dfrac{1}{(N+1)}, & \psi = 1. \end{cases}$$

Следует отметить, что выполнение условия стационарности $\psi = \dfrac{\lambda}{\mu} < 1$ для данной СМО не обязательно, поскольку число допускаемых в обслуживающую систему заявок контролируется путем введения ограничения на длину очереди (которая не может превышать N - 1), а не соотношением между интенсивностями входного потока, т. е. не отношением $\psi = \dfrac{\lambda}{\mu}$.

Определим характеристики одноканальной СМО с ожиданием и ограниченной длиной очереди, равной (N - 1):

- вероятность отказа в обслуживании заявки:

$$P_{отк} = P_N = \begin{cases} \left(\dfrac{1-\psi}{1-\psi^{N+1}}\right)\cdot \psi^n, & \psi \neq 1, \\ \dfrac{1}{(N+1)}, & \psi = 1. \end{cases}$$

- относительная пропускная способность системы:

$$q = 1 - P_{отк} = \begin{cases} 1 - \left(\dfrac{1-\psi}{1-\psi^{N+1}}\right)\cdot \psi^n, & \psi \neq 1, \\ 1 - \dfrac{1}{(N+1)}, & \psi = 1. \end{cases}$$

- абсолютная пропускная способность:

$$A = q \cdot \lambda$$

- среднее число находящихся в системе заявок:

$$L_S = \sum_{n=0}^{N} n \cdot P_n = \begin{cases} \dfrac{\psi \cdot \left[1 - (N+1) \cdot \psi^N + N \cdot \psi^{N+1}\right]}{(1-\psi) \cdot \left(1 - \psi^{N+1}\right)}, & \psi \neq 1 \\ N/2, \quad \psi = 1 \end{cases}$$

- среднее время пребывания заявки в системе:

$$W_S = \frac{L_S}{\lambda(1 - P_N)}$$

- средняя продолжительность пребывания клиента (заявки) в очереди:

$$W_q = W_S - \frac{1}{\mu}$$

- среднее число заявок (клиентов) в очереди (длина очереди):

$$L_q = \lambda(1 - P_N)W_q$$

Решение одноканальной модели массового обслуживания с отказами

Пусть одноканальная СМО с отказами представляет собой один пост ежедневного обслуживания (ЕО) для мойки автомобилей. Заявка - автомобиль, прибывший в момент, когда пост занят, - получает отказ в обслуживании. Интенсивность потока автомобилей λ= 1,0 (автомобиль в час). Средняя продолжительность обслуживания - 1,8 часа. Поток автомобилей и поток обслуживании являются простейшими.

Требуется определить в установившемся режиме предельные значения:

- относительной пропускной способности q;

- абсолютной пропускной способности А;

- вероятности отказа *Ротк* ;

Сравните фактическую пропускную способность СМО с номинальной, которая была бы, если бы каждый автомобиль обслуживался точно 1,8 часа и автомобили следовали один за другим без перерыва.

Решение

1. Определим интенсивность потока обслуживания:

$$\mu = \frac{1}{\bar{t}_{об}} = \frac{1}{1{,}8} = 0{,}555$$

.

2. Вычислим относительную пропускную способность:

$$q = \frac{\mu}{\mu + \lambda} = \frac{0,555}{1 + 0,555} = 0,356$$
.

Величина q означает, что в установившемся режиме система будет обслуживать примерно 35% прибывающих на пост ЕО автомобилей.

3. Абсолютную пропускную способность определим по формуле:

$$A = \lambda \cdot q = 1 \cdot 0,356 = 0,356$$
.

Это означает, что система (пост ЕО) способна осуществить в среднем 0,356 обслуживания автомобилей в час.

4. Вероятность отказа:

$P_{отк} = 1 - q = 1 - 0,356 = 0,644$.массовый обслуживание отказ модель

Это означает, что около 65% прибывших автомобилей на пост ЕО получат отказ в обслуживании.

5. Определим номинальную пропускную способность системы:

$$A_{ном} = \frac{1}{\bar{t}_{обсл}} = \frac{1}{0,8} = 0,555$$
(автомобилей в час).

Оказывается, что *Аном* в 1,5 раза $\left(\frac{0,555}{0,356} \approx 1,5\right)$ больше, чем фактическая пропускная способность, вычисленная с учетом случайного характера потока заявок и времени обслуживания.

3.3 Практическое занятие № 3

Тема: «Характеристика систем массового обслуживания»

1. Работа каждого канала характеризуется временем, которое затрачивается на обслуживание одной заявки. В общем случае это время является случайным. Для простейших пуассоновских систем время обслуживания распределено по показательному закону с параметром μ. Это эквивалентно тому, что на выходе непрерывно занятого канала будет простейший поток обслуженных заявок с параметром μ. Как правило, все каналы имеют одинаковую интенсивность обслуживания μ. В этом случае нет необходимости различать каналы (первый, второй, и т. д.). Помимо этого, будем считать, что заявка может обслуживаться любым из n каналов, т. е. любой из n каналов «доступен» для заявки.

2. Следующим важным параметром СМО является интенсивность (плотность) потока заявок λ. Здесь уместно напомнить, что заявки различаются лишь моментом

поступления на обслуживание, а интенсивность потока заявок λ определяется через средний интеграл между поступлениями двух заявок:

$$\lambda = 1 / t_\lambda$$

Каждый канал обслуживания обеспечивается g пусковыми установками, а каждая пусковая установка производит в среднем $\overline{\mu}$ выстрелов в минуту. Тогда эффективная скорострельность одного канала определится так:

$$\mu = g\overline{\mu}p,$$

где p – вероятность поражения цели одной выпущенной по ней ракетой.

Среднее время пребывания цели в зоне обстрела:

$$M[T_з] = \overline{t_з} = \frac{a}{V},$$

где V – скорость полета самолета, при условии, что она обстреливается.

Далее, в силу того, что мы рассматриваем только пуассоновские системы, будем считать, что время пребывания самолёта в зоне обстрела $T_з$ является показательным с параметром η:

$$\eta = \frac{1}{\bar{t}_з} = \frac{V}{a}.$$

Тогда время занятости канала $T_{об}$ будет также подчинено показательному закону с параметром

$$\mu^* = \frac{\mu}{1+\mu\bar{t}_и} + \eta.$$

Таким образом, поток освобождений канала ПВО, определяемый этим выражением, имеет интенсивность μ^*

$$\mu^* = \mu + \eta.$$

Таким образом, видно, что поток освобождений канала слагается из двух потоков: потока поражающих выстрелов с параметром μ и потока уходов непоражённых самолётов из зоны обстрела с параметром η. Другими словами канал освобождается либо по причине поражения самолёта, либо по причине выхода самолёта из зоны обстрела непоражённым. Если время передачи информации велико по сравнению со временем пребывания цели в зоне обстрела (типично для системы ПВО, у которой время $T_з$ мало), то

$$\mu^* = \eta$$

При анализе работы системы ПВО необходимо знать характеристики налета. Будем считать, что налетающие самолеты образуют пуассоновский поток с интенсивностью λ, который определяется так:

$$\lambda = \frac{V}{I},$$

где I – средний линейный интервал между самолетами.

В нашем случае время передачи информации мало и параметр освобождения канала определяется по формуле $\mu^* = \mu + \eta$

3.4 Практическое занятие № 4

Тема: «Игровые методы обоснования решений»

Имеется два игрока А и В: игрок А прячется, а В его ищет. В распоряжении А имеется два убежища (I и II), любое из которых он может выбрать по своему усмотрению. Условия игры таковы: если В найдет А в том убежище, где А спрятался, то А платит ему штраф 1 руб ; если В не найдет А (т. е. будет искать в другом убежище), то он сам должен заплатить А такой же штраф. Требуется построить платежную матрицу.

Решение. Игра состоит всего из двух ходов, оба — личные. У нас (А) две стратегии:

А1— прятаться в убежище 1,

А2— прятаться в убежище II.

У противника (В) тоже две стратегии:

В1 — искать в убежище I,

В2— искать в убежище II.

Перед нами — игра 2×2. Ее матрица имеет вид:

Bj Ai	B1	B2
A1	-1	1
A2	1	-1

На примере этой игры, как она ни элементарна, можно уяснить себе некоторые важные идеи теории игр. Предположим сначала, что данная игра выполняется только один раз (играется единственная «партия»). Тогда, очевидно, нет смысла говорить о преимуществах тех или других стратегий — каждый из игроков может с равным основанием принять любую из них. Однако при многократном повторении игры положение меняется.

Действительно, допустим, что мы (игрок А) выбрали какую-то стратегию (скажем, А1) и придерживаемся ее. Тогда, уже по результатам первых нескольких партий, противник догадается о нашей стратегии, начнет всегда искать в убежище I и выигрывать. То же будет, если мы выберем стратегию А2 . Нам явно невыгодно придерживаться одной какой-то стратегии; чтобы не оказаться в проигрыше, мы должны чередовать их. Однако, если мы будем чередовать убежища I и II в какой-то определенной последовательности (скажем, через одну партию), противник тоже догадается об этом и ответит наихудшим для нас образом. Очевидно, надежным способом, гарантирующим нас от верного проигрыша, будет такая организация выбора в каждой партии, когда мы сами его наперед не знаем. Например, можно бросить монету, и, если выпадет герб, выбрать убежище I, а если решка II.

Печальное положение, в котором оказался игрок А (чтобы не проигрывать, выбирать убежище случайным образом), очевидно, присуще не только ему, но и его противнику В, для которого справедливы все вышеприведенные рассуждения. Оптимальной стратегией каждого оказывается «смешанная» стратегия, в которой две возможные стратегии игрока чередуются случайным образом, с одинаковыми вероятностями.

Таким образом, мы путем интуитивных рассуждений подошли к одному из существенных понятий теории игр — к понятию смешанной стратегии — т. е. такой, в которой отдельные «чистые» стратегии чередуются случайным образом с какими-то вероятностями. В данном примере из соображений симметрии ясно, что стратегии А1 и А2 должны применяться с одинаковыми вероятностями; в более сложных примерах решение может быть далеко не тривиальным.

3.5 Практическое занятие № 5

Тема: «Методы решения конечных игр»

Перед тем, как решать игру m X n, нужно, прежде всего, попытаться ее упростить, избавившись от лишних стратегий. Это делается подобно тому, как мы когда-то отбрасывали заведомо невыгодные решения в § 6. Введем понятие «доминирования». Стратегия A_i игрока A называется доминирующей над стратегией A_k, если в строке A_i

стоят выигрыши не меньшие, чем в соответствующих клетках строки $A_к$, и из них по крайней мере один действительно больше, чем в соответствующей клетке строки $A_к$. Если все выигрыши строки A_i равны соответствующим выигрышам строки

A_k, то стратегия A_i называется дублирующей стратегию $A_к$. Аналогично определяются доминирование и дублирование для стратегий игрока *B:* доминирующей называется та его стратегия, при которой везде стоят выигрыши не большие, чем в соответствующих клетках другой, и по крайней мере один из них действительно меньше; дублирование означает полное повторение одного столбца другим. Естественно, что если для какой-то стратегии есть доминирующая, то эту стратегию можно отбросить; также отбрасываются и дублирующие стратегии.

Поясним сказанное примером. Пусть игра 5 X 5 задана матрицей таблицы 27.1.

Таблица 27.1

Прежде всего заметим, что в ней стратегия A_5 дублирует стратегию A_2, поэтому любую из них можно отбросить. Отбрасывая A_5, замечаем, что в строке *A* все выигрыши больше (или равны) соответствующим выигрышам строки A_4, значит A_1 доминирует над A_4. Отбросим A_4 и получим матрицу 3 X 5 (таблица 27.2).

Но это еще не все! Приглядевшись к таблице 27.2, замечаем, что в ней некоторые стратегии игрока *B* доминируют над другими: например, B_3 над B_4 *и* B_5, а B_1 — над стратегией B_2 (не забудьте, что *B* стремится отдать поменьше!). Отбрасывая столбцы B_2, B_4, B_5, получаем игру 3X2 (таблица 27.3).

Таблица 27.2

Наконец, в таблице 27.3 строка A_3 дублирует A_1, поэтому ее можно отбросить. Окончательно получим игру 2X2 (таблица 27.4).

Таблица 27.3

Таблица 27.4

Эту игру, как ни старайся, уже не упростишь. Приходится решать. Попутно заметим, что, отбрасывая лишние (дублирующие и заведомо невыгодные) стратегии в игре с седловой точкой, мы придем к решению в чистых стратегиях. Но лучше сразу проверить, не обладает ли игра седловой точкой — это проще, чем сравнивать почленно все строки и все столбцы.

В руководствах по теории игр обычно останавливаются на решении простейших игр 2 X 2, 2 X *n* и *m* X 2, которое допускает геометрическую интерпретацию, но мы этого делать не будем — сразу возьмем «быка за рога» и покажем, как можно решить любую игру *m* X *n*,

Пусть имеется игра *m* X *n* без седловой точки с матрицей (a_{ij}) (см. таблицу 27.5).

Таблица 27.5

Допустим, что все выигрыши a_{ij} положительны (этого всегда можно добиться, прибавляя ко всем членам матрицы достаточно большое число M; от этого цена игры увеличится на M, а решение S^*_A, S^*_B В не изменится). Если все a_{ij} положительны, то конечно, и цена игры, т. е. средний выигрыш при оптимальной стратегии, тоже положителен: $v > 0$.

Мы хотим найти решение игры, т. е. две оптимальные смешанные стратегии

дающие каждой стороне максимально возможный для нее средний выигрыш (минимальный проигрыш).

Найдем сначала S^*_A. Мы знаем, что если один из игроков (в данном случав это A) применяет свою оптимальную стратегию, то другой (B) не может улучшить свое положение, отступая от своей. Заставим противника (B) отступать от своей оптимальной стратегии, пользуясь чистыми стратегиями B_1, B_2, ..., B_n (а мы тем временем упорно держимся стратегии S^*_A). В любом случае наш выигрыш будет не меньше, чем v:

Разделим неравенства (27.2) на положительную величину v и введем обозначения:

Тогда условия (27.2) примут вид

где x_1, x_2, ..., x_m — неотрицательные переменные. В силу (27.3) и того, что $p_1 + p_2 +... + p_m = 1$, переменные x_1, x_2, ..., x_m удовлетворяют условию

Но v есть не что иное, как наш г а р а н т и р о в а н ный выигрыш; естественно, мы хотим сделать его максимальным, а значит, величину $1/v$ -- минимальной.

Таким образом, задача решения игры свелась к математической задаче: найти неотрицательные значения переменных x_1, x_2, ..., x_m такие, чтобы они удовлетворяли линейным ограничениям-неравенствам (27.4) и обращали в минимум линейную функцию этих переменных:

«Ба! — скажет читатель,— что-то знакомое!» И точно — перед нами не что иное, как *задача линейного программирования.* Таким образом, задача решения игры m Х n свелась к задаче линейного программирования с n ограничениями-неравенствами и m переменными. Зная x_1, x_2, ..., x_m, можно по формулам (27.3) найти p_1, p_2, ..., p_m и, значит, оптимальную стратегию S^*_A и цену игры v.

Оптимальная стратегия игрока B находится совершенно аналогично, с той разницей, что B стремится минизировать, а не максимизировать выигрыш, а значит, обратить не в минимум, а в максимум величину $1/v$, а в

ограничениях-неравенствах вместо знаков $\geq$ будут стоять $\leq$. Пара задач линейного программирования, по которой находятся оптимальные стратегии (S^*_A, S^*_B)В , называется парой двойственных з а д а ч линейного п р о гр а мм и р о в а ни я (доказано, что максимум линейной функции в одной из них равен минимуму линейной функции в другой, так что все в порядке — разных значений цены игры мы не получим).

Таким образом, решение игры m Х n эквивалентно решению задачи линейного программирования. Нужно заметить, что и наоборот,— для любой задачи линейного программирования может быть построена эквивалентная ей задача теории игр (на том, как это делается, мы останавливаться не будем). Эта связь задач теории игр с задачами линейного программирования оказывается полезной не только для теории игр, но и для линейного программирования. Дело в том, что существуют приближенные численные методы решения игр, которые в некоторых случаях (при большой размерности задачи) оказываются проще, чем «классические» методы линейного программирования.

Опишем один из самых простых численных методов решения игр — так называемый метод и т е р а ц и й (иначе — метод Брауна — Робинсон). Идея его в следующем. Разыгрывается «мысленный эксперимент», в котором стороны A и B поочередно применяют друг против друга свои стратегии, стремясь выиграть побольше (проиграть поменьше). Эксперимент состоит из ряда «партий» игры. Начинается он с того, что один из игроков (скажем, A) выбирает произвольно одну из своих стратегий A_t. Противник (B) отвечает ему той из своих стратегий B), которая хуже всего для A, т. е. обращает его выигрыш при стратегии A_i в минимум. Дальше снова очередь A — он отвечает B той своей стратегией A_k, которая дает максимальный выигрыш при стратегии B_j противника. Дальше — снова очередь противника. Он отвечает нам той своей стратегией, которая является наихудшей не для последней, примененной нами, стратегии $A_к$, а для смешанной стратегии , в которой до сих пор примененные стратегии A_i, $A_к$ встречаются с равными вероятностями. И так далее: на каждом шаге итерационного процесса каждый игрок отвечает на очередной ход другого *той своей стратегией, которая является оптимальной для него относительно смешанной стратегии другого, в* *которую все примененные до сих пор стратегии входят пропорционально частотам их применения.* Вместо того чтобы вычислять каждый раз средний выигрыш, можно пользоваться просто «накопленным» за предыдущие ходы выигрышем и выбирать ту свою стратегию, при которой этот накопленный выигрыш максимален (минимален). Доказано, что такой метод сходится: при увеличении числа «партий» средний выигрыш на одну партию будет стремиться к цене игры, а частоты применения стратегий — к их вероятностям в оптимальных смешанных стратегиях игроков.

Впрочем, лучше всего можно понять итерационный метод на конкретном примере. Продемонстрируем его на примере игры 3 X 3 предыдущего параграфа (таблица 26.5). Чтобы не иметь дела с отрицательными числами, прибавим к элементам матрицы таблицы 26.5 число 5 (см. таблицу 27.6); при этом цена игры увеличится на 5, а решение S^*_A, S^*_B В не изменится.

Начнем с произвольно выбранной стратегии игрока *А,*— например, со стратегии A_3. В таблице 27.7 приведены первые 15 шагов итерационного процесса по методу

Брауна — Робинсон (читатель может самостоятельно продолжить расчеты). В первом столбце дан номер пар, тии (пары выборов) *к,* во втором — номер i выбранной в данной партии стратегии игрока *А*. В последующих трех столбцах — «накопленный выигрыш» за первые k партий при тех стратегиях , которые применяли игроки в предыдущих партиях и при стратегиях B_1, B_2, B_3 игрока B в данной партии (получается прибавлением элементов соответствующей строки к тому, что было строкой выше). Из этих накопленных выигрышей в таблице 27.7 подчеркнут минимальный (если их несколько, подчеркиваются все). Подчеркнутое число определяет ответный выбор игрока B в данной партии — он выбирает ту стратегию, которая

Таблица 27.6

соответствует подчеркнутому числу **(**если их несколько, берется любая). Таким образом определяется номер j оптимальной (в данной партии) стратегии B (ставится в следующем столбце). В последующих

Т а б л и ц а 27.7

трех столбцах дается накопленный выигрыш за $к$ партий соответственно при стратегиях A_1, A_2, A_3 игрока A (получается прибавлением элементов столбца B_jВ к тому, что было строкой выше). Из этих значений в таблице 27.7 «надчеркнуто» максимальное; оно определяет выбор стратегии игрока A в следующей партии (строкой ниже). В последних трех столбцах таблицы 27.7 даны: v — нижняя оценка цены игры, равная минимальному накопленному выигрышу, деленному на число партий k; v — верхняя оценка цены игры, равная максимальному накопленному выигрышу, деленному на k; v^* — среднее арифметическое между ними (оно служит лучшей, чем нижняя и верхняя, приближенной оценкой цены игры).

Как видно, величина v^* незначительно колеблется около цены игры $v = 5$ (цена исходной игры была 0, но мы прибавили к элементам матрицы по 5). Подсчитаем по таблице 27.7 частоты p_1, p_2, p_3, q_1, q_2, q_3 стратегий игроков. Получим:

$\tilde{p}_1 = 4/15 \approx 0{,}266$, $\tilde{p}_2 = 7/15 \approx 0{,}468$, $\tilde{p}_3 = 4/15 \approx 0{,}266$, $\tilde{q}_1 = 2/15 \approx 0{,}133$, $\tilde{q}_2 = 8/15 \approx 0{,}534$, $\tilde{q}_3 = 5/15 \approx 0{,}333$,

что не так уж сильно отличается от вероятностей p_1, p_2, p_3, q_1, q_2, q_3, равных, как мы указывали раньше, для первой, второй и третьей стратегий соответственно 1/4 = 0,25, 1/2 = 0,50, 1/4 = 0,25. Такие сравнительно хорошие приближения мы получили уже при 15 итерациях — это обнадеживает! К сожалению, дальше процесс приближений будет идти не так резво. Сходимость метода Брауна — Робинсон, как показывает опыт, очень медленная, Существуют способы, позволяющие как бы «подхлестнуть» еле плетущийся процесс, но мы на них останавливаться не будем.

Очень важным преимуществом итерационного метода решения игр является то, что его трудоемкость сравнительно медленно возрастает с увеличением размерности игры m X n, тогда как трудоемкость метода линейного программирования растет при увеличении размерности задачи гораздо быстрее.

* * *

Таким образом, читатель получил некоторое представление о теории антагонистических игр и о методах решения матричных игр.

Скажем несколько критических слов по поводу этой теории и ее практической значимости. В свое время, когда теория игр еще только появилась, на нее возлагались большие надежды в смысле выбора решений для конфликтных ситуаций. Эти надежды оправдались лишь в малой степени.

Прежде всего, на практике не так уж часто встречаются строго антагонистические конфликты — разве только в настоящих «играх» (шашки, шахматы, карты и т. п.). Вне этих искусственных ситуаций, где одна сторона стремится во что бы то ни стало обратить выигрыш в максимум, а другая — в минимум, такие конфликты почти не встречаются. Казалось бы, где, как не в области боевых действий должна была бы с успехом применяться теория игр? Ведь здесь мы встречаемся с самыми «свирепыми» антагонизмами, с самой резкой противоположностью интересов! Но оказывается, что конфликтные ситуации в этой области сравнительно редко удается свести к парным играм с нулевой суммой. Схема строгого антагонизма применима, как правило, только к операциям малого масштаба, ограниченным по значению. Например, сторона A — группа самолетов, налетающих на объект, сторона B — средства противовоздушной обороны объекта; первая стремится максимизировать вероятность уничтожения объекта, вторая — ее минимизировать. Здесь схема парной игры с нулевой суммой может найти применение. Но возьмем чуть более сложный пример: две группировки боевых единиц (типа танков, самолетов, кораблей) ведут бой. Каждая сторона стремится поразить как можно больше боевых единиц противника. В этом примере антагонизм ситуации теряет свою чистоту: она уже не сводится к парной игре с нулевой суммой. Если цели участников конфликта не прямо противоположны, а просто не совпадают, математическая модель становится много сложнее: мы уже не можем интересоваться выигрышем только одной стороны; возникает так называемая «биматричная игра», где каждый из участников стремится максимизировать свой выигрыш, а не просто минимизировать выигрыш противника. Теория таких игр гораздо сложнее, чем теория антагонистических игр, а главное, из этой теории не удается получить четких рекомендаций по оптимальному образу действий сторон .

Вторoе критическое замечание будет касаться понятия «смешанных стратегий». Если речь идет о многократно повторяемой ситуации, в которой каждая сторона может легко (без дополнительных затрат) варьировать свое поведение от случая к случаю, оптимальные смешанные стратегии в самом деле могут повысить средний выигрыш. Но бывают ситуации, когда решение надо принять одно-единственное (например, выбрать план строительства системы оборонительных сооружений). Разумно ли будет «передоверить свой выбор случаю»,—грубо говоря, бросить монету, и если выпал герб, выбрать первый вариант плана, а если решка — второй? Вряд ли найдется такой руководитель, который в сложной и ответственной ситуации решится делать выбор случайным образом, хотя бы это и вытекало из принципов теории игр!

Наконец, последнее соображение: в теории игр считается, что каждому игроку известны все возможные стратегии противника, неизвестно лишь то, какой именно из них он воспользуется в данной партии игры. В реальном конфликте это обычно не так: перечень возможных стратегий противника как раз неизвестен, и наилучшим решением в конфликтной ситуации нередко будет именно *выйти за пределы известных противнику стратегий,* «ошарашить» его чем-то совершенно новым, непредвиденным!

Как видно из вышеизложенного, теория игр в качестве основы для выбора решения (даже в остроконфликтной ситуации) имеет много слабых мест. Значит ли это, что ее не нужно изучать, что она вовсе не нужна в исследовании операций? Нет, не значит. Теория игр ценна прежде всего самой постановкой задач, которая учит, выбирая решение в конфликтной ситуации, не забывать о том, что противник тоже мыслит, и учитывать его возможные хитрости и уловки. Пусть рекомендации, вытекающие из игрового подхода, не всегда определенны и не всегда осуществимы --все же полезно, выбирая решение, ориентироваться, в числе других, и на игровую модель. Не надо только выводы, вытекающие из этой модели, считать окончательными и непререкаемыми .

3.6 Практическое занятие № 6

Тема: «Критерии принятия решения в условиях неопределенности»

Теория статистических решений может быть истолкована как теория поиска оптимального недетерминированного поведения в условиях неопределенности. Согласно А.Вальду, поведение считается оптимальным, если оно минимизирует риск в последовательных экспериментах, т.е. математическое ожидание убытков статистического эксперимента.

Считаю необходимым рассмотреть четыре критерия принятия решений в условиях неопределенности, когда никакие вероятностные характеристики не известны.

- критерий Лапласа,
- минимаксный критерий,
- критерий Сэвиджа,
- критерий Гурвица.

Основное различие между этими критериями определяется стратегией лица, принимающего решения. Критерий Лапласа основан на более оптимистичных предположениях, чем минимаксный критерий. Критерий Гурвица можно использовать при различных подходах – от наиболее оптимистичного до наиболее пессимистичного. Все эти критерии отражают субъективную оценку ситуации, в которой приходится принимать решение. При этом не существует общих правил применимости того или иного критерия, так как поведение лица, принимающего решение в условиях неопределенности, является наиболее важным фактором при выборе подходящего критерия.

Перечисленные критерии базируются на том, что лицу, принимающему решение, не противостоит разумный противник. В случае, когда в роли противника выступает природа, нет оснований предполагать, что она стремится причинить вред лицу, принимающему решение.

При наличии разумного противника, интересы которого противоречат интересам лица, принимающего решения (например, в военных действиях противоборствующие армии являются разумными противниками), для построения подходящего критерия требуется специальный подход. Эти вопросы рассматриваются в теории игр.

Данные, необходимые для принятия решений в условиях неопределенности, задаются в форме матрицы, строки которой соответствуют действиям, а столбцы - возможным состояниям системы.

Каждому действию и каждому возможному состоянию системы соответствует результат (исход), определяющий выигрыш (или потери) при выборе данного действия и реализации данного состояния.

Пусть ai (i=1,2, ... , m)

и q j представляет возможное состояние j (j=1,2, ... ,n),

n (ai , q j) - описывает соответствующий результат.

В общем случае n (ai , q j) может быть непрерывной функцией ai и q j .

В дискретном случае указанные данные представляются в форме матрицы.

	q 1	q 2	...	q n
a1	n (a1 ,q 1)	n (a1 ,q 2)	...	n (a1 ,q n)
a2	n (a2 ,q 1)	n (a2 ,q 2)	...	n (a2 ,q n)
...	...	...	...	...
am	n (am ,q 1)	n (am ,q 2)	...	n (am ,q n)

Критерий Лапласа

Этот критерий опирается на известный принцип недостаточного обоснования. Поскольку вероятности состояний q 1, q 2, ... ,q n не известны, необходимая информация для вывода, что эти вероятности различны, отсутствует. В противном случае можно было бы определить эти вероятности и ситуацию уже не следовало рассматривать как принятие решения в условиях неопределенности. Так как принцип недостаточного обоснования утверждает противоположное, то состояния q 1, q 2, ...,q n имеют равные вероятности. Если согласиться с приведенными доводами, то исходную задачу можно рассматривать как задачу принятия решений в условиях риска, когда выбирается действие ai , дающее ожидаемый выигрыш.

Другими словами, находится действие ai* , соответствующее

$$\max_{a_i}\left\{\frac{1}{n}\sum_{j=1}^{n} x(a_i, \theta_j)\right\}$$

$\frac{1}{n}$ - вероятность реализации состояния q j (j=1,2, ... ,n),

Пример. Одно из предприятий должно определить уровень предложения услуг так, чтобы удовлетворить потребности клиентов в течение предстоящих праздников. Точное число клиентов не известно, но ожидается, что оно может принять одно из четырех значений: 200, 250, 300 или 350 клиентов. Для каждого из этих возможных значений существует наилучший уровень предложения (с точки зрения возможных затрат). Отклонения от этих уровней приводят к дополнительным затратам либо из-за превышения предложения над спросом, либо из-за неполного удовлетворения спроса.

В таблице приведены потери в тысячах долларов.

Клиенты

Уровень предложения

	q 1	q 2	q 3	q 4
a1	5	10	18	25
a2	8	7	8	23
a3	21	18	12	21
a4	30	22	19	15

Принцип Лапласа предполагает, что q 1, q 2, q 3, q 4 равновероятны.

Следовательно, P{q =q j } =1/4, j= 1, 2, 3, 4, и ожидаемые потери при различных действиях a1, a2, a3, a4 составляют

E{a1}= (1/4)(5+10+18+25)=14,5

E{a2}= (1/4)(8+7+8+23)=11,5

E{a3}= (1/4)(21+18+12+21)=18,0

E{a4}= (1/4)(30+22+19+15)=21,5

Таким образом, наилучшим уровнем предложения в соответствии с критерием Лапласа будет a2.

Минимаксный (максиминный) критерий

Является наиболее осторожным, поскольку основывается на выборе наилучшей из наихудших возможностей. Если результат n (ai , q j) представляет потери лица, принимающего решение, для действия ai наибольшие потери независимо от возможного состояния q j будут равны

$$\max_{\theta_j}\left\{\nu\left(a_i,\theta_j\right)\right\}$$

По **минимаксному** критерию должно выбираться действие a_i, дающее

$$\min_{a_i}\max_{\theta_j}\left\{\nu\left(a_i,\theta_j\right)\right\}$$

Аналогично в том случае, когда ν $(a_i\ ,\ \theta_j)$ представляет выигрыш, согласно минимаксному критерию, выбирается действие a_i , дающее

$$\max_{a_i}\min_{\theta_j}\left\{\nu\left(a_i,\theta_j\right)\right\}$$

В этом случае критерий называется максиминным.

Пример. Рассмотрим предыдущий пример. Так как n (ai , q j) представляют потери, применим минимаксный критерий. Результаты вычислений представим в виде следующей таблицы.

	q 1	q 2	q 3	q 4	$\max\limits_{\theta_j}\left[s(a_i, \theta_j) \right]$
a1	5	10	18	25	25
a2	8	7	8	23	23
a3	21	18	12	21	21
a4	30	22	19	15	30

Минимаксной стратегией будет a3 .

Подходы к учету неопределенности при описании рисков. В теории принятия решений в настоящее время при компьютерном и математическом моделировании для описания неопределенностей чаще всего используют такие математические средства, как:

- вероятностно-статистические методы,

- методы статистики нечисловых данных, в том числе интервальной статистики и интервальной математики, а также методы теории нечеткости,

- методы теории конфликтов (теории игр).

Они применяются в имитационных, эконометрических, экономико-математических моделях, реализованных обычно в виде программных продуктов.

Некоторые виды неопределенностей связаны с безразличными к организации силами - природными (погодные условия) или общественными (смена правительства). Если явление достаточно часто повторяется, то его естественно описывать в вероятностных терминах. Так, прогноз урожайности зерновых вполне естественно вести в вероятностных терминах. Если событие единично, то вероятностное описание вызывает внутренний протест, поскольку частотная интерпретация вероятности невозможна. Так, для описания неопределенности, связанной с исходами выборов или со сменой правительства, лучше использовать методы теории нечеткости, в частности, интервальной математики (интервал – удобный частный случай описания нечеткого множества). Наконец, если неопределенность связана с активными действиями соперников или партнеров, целесообразно применять методы анализа конфликтных ситуаций, т.е. методы теории игр, прежде всего антагонистических игр, но иногда полезны и более новые методы кооперативных игр, нацеленных на получение устойчивого компромисса.

Иногда под уменьшением риска понимают уменьшение дисперсии случайной величины, поскольку при этом уменьшается неопределенность. В теории принятия решений риск - это плата за принятие решения, отличного от оптимального, он обычно выражается как математическое ожидание. В экономике плата измеряется обычно в денежных единицах, т.е. в виде финансового потока (потока платежей и поступлений) в условиях неопределенности.

Критерий Сэвиджа

Этот критерий характеризуется крайней осторожной (пессимистической) позицией к возможным потерям из-за отсутствия достоверных сведений о том, какая из ситуаций, влияющих на экономический результат, будет иметь место в конкретном случае. Реализуется применительно к матрице рисков и потерь.

Матрица потерь строится следующим образом:

1.Находим наибольшее значение по каждому случайному событию Qi

2. Выписываем их в качестве утопических точек отдельно

3.Вычитаем из каждой такой утопической точки соответствующие этому случайному события Xi (пример: для Q1: Xy-X1,Xy-X2,Xy-X3.....).

4.Получаем новую матрицу потерь.

В рамках такого подхода функция, задающая семейство «линий уровня» определяется равенством:

$F(u,v,......,z)= \max(a_y-u,\ a_y-v,......,\ a_y-z)$

Целевая функция критерия:

Zs=min(Ki), где Ki=max(Lij), Lij=max(Aij)-Ay, где (Lij) – матрица потерь

i – вариант возможного решения ЛПР

j – вариант возможной ситуации

Aij – доход ЛПР, если будет принято решение i, а ситуация сложится j

A = (Aij) – матрица полезностей.

(Lij) – соответствующая матрица рисков или потерь

Критерий Гурвица

Критерий Гурвица – это взвешенная позиция “пессимизма-оптимизма”.

При C =1 - критерий Гурвица просто соответствует Максиминному критерию.

Составные критерия принятия решений в условиях неопределенности.

Шаг А: требования к допустимому риску.

Вот на этом шаге уточняется критический уровень дохода(или потерь), приемлемый для ЛПР в конкретной ситуации. За основу бреется опорное значение для выбранного опорного критерия. После задается допустимое для ЛПР максимально возможное отклонение Eдоп>0 от опорного значения(в худшую сторону).

Шаг Б: блокировка решений с недопустимом риском.

Вот на этом шаге удаляются из исходной матрицы все решения, который не подходят требованиям ЛПР, которые предъявляются к допустимому риску применительно к анализируемой ситуации.

Шаг В: требования к компенсации за риск.

Этот шаг уточняет требования к анализируемым решениям, для которых баланс между риском потерь(при -) и компенсации(при +) является приемлемым для ЛПР.

Шаг Г: блокировка решений с недостаточной компенсацией риска.

Вот на этом шаге из матрицы полезностей(которая будет получена после шага Б) удаляются все решения, которые не соответствуют требованиям ЛПР.

Шаг Д: выбор оптимального решения.

И наконец, на этом шаге для оставшейся «урезанной» матрицы находится оптимальное решение по заранее оговоренном критерию. Это найденное решение и будит являться оптимальным выбором для соответствующего составного критерия.

Последствия решений менеджера, экономиста, инженера проявятся в будущем. А будущее неизвестно. Мы обречены принимать решения в условиях неопределенности. Мы всегда рискуем, поскольку нельзя исключить возможность нежелательных событий. Но можно сократить вероятность их появления. Для этого необходимо спрогнозировать дальнейшее развитие событий, в частности, последствия принимаемых решений.

Задача №1.

Предприятие выпускает два вида продукции: А и В. При этом используются ресурсы: Rl, R2 и R3. Нормы расхода на ресурсы составляют соответственно:

R1: a1, a2

R2: b1,b2

R3: c1, c2

Рыночная цена продукции А составляет-P1, продукции В-Р2. Необходимо принять решение относительно плана выпуска продукции обеспечивающего максимальный доход. Оценить устойчивость выбранного решения относительно колебания цен на продукцию. Объемы ресурсов: Rl -Vl, R2-V2, R3-V3

Вариант	al	a2	bl	Ь2	cl	c2	P1	P2	VI	V2	V3
12	3	5	2	1	4	6	3	2	30	20	48

Обозначим x_1 - количество продукции А, x_2 - Количество продукции В.

Найти X=(x_1, x_2), удовлетворяющие системе

$3x_1+5x_2 \leq 30$ -количество ресурса R_1

$2x_1+x_2 \leq 20$ -количество ресурса R_2

$4x_1+6x_2 \leq 48$ - количество ресурса R_3

и условию $x_j \geq 0$

при котором функция дохода принимает максимальное значение.

$V = P1\,x_1 + P2\,x_2 = 3\,x_1 + 2\,x_2 \rightarrow \max$

Формулировка задачи.

Графический метод.

Построим ОДЗ x_1 и x_2

Неравенства $x_1 \geq 0$, $x_2 \geq 0$ задают первый квадрант координатной плоскости.

Неравенство $3x_1+5x_2$□□□ задает полуплоскость, расположенную под прямой $3x_1+5x_2=30$, включая эту прямую.

Неравенство $2x_1+x_2$□20 задает полуплоскость, расположенную под прямой $2x_1+x_2=20$, включая эту прямую.

Неравенство $4x_1+6x_2$□48 задает полуплоскость, расположенную под прямой $4x_1+6x_2=48$, включая эту прямую.

Таким образом, получаем, что множество точек, удовлетворяющее всем неравенствам, Область ОАВС.

Построим вектор N{3;2}. Его проекция на ось OX_1 равна 3, на ось OX_2 2.

Поскольку необходимо найти максимум функции *V*, будем перемещать прямую *l*, перпендикулярно вектору *Н,* от начала к концу вектора *Н,* т.е. в направлении возрастания функции *V*. Перейдя в точку В, прямая *l* окажется на выходе из многоугольной области ОАВС. Точка В – (крайняя) последняя точка области при движении в направлении вектора *Н*, поэтому значение функции *V* в этой точке будет наибольшим по сравнению с ее значениями в других точках области.

Поскольку точка В – точка пересечения первой и второй прямой, то ее координаты можно найти, решая систему уравнений:

□□□x1 +5x2 □□□□

□□□

□□ □□□□□□ x_1 □ x_2 □□□□

Выразим из второго уравнения x_2 :

x_2 = 20-2x_1

И подставим в первое уравнение

3x_1+5(20-2x_1) = 30

Откуда x_1 = 10

Подставив x_1 в выражение для x_2 , получим $x2$ = 0

Таким образом оптимальное решение – точка В (10,0)

Оценим устойчивость выбранного решения относительно колебания цен на продукцию.

Функция V=3x_1+2x_2 достигает максимального значения в угловой точке В. При изменения коэффициентов целевой функции $V = P_1x_1 + P_2x_2$ точка В останется точкой оптимального решения до тех пор, пока угол наклона прямой l будет лежать между углами наклона двух прямых, пересечением которых является точка В. Этими прямыми являются $3x_1 + 1x_2 = 120$ (ограничение на ресурс R1) и $5x_1 + 1x_2 = 150$ (ограничение на ресурс R2).

Алгебраически записывается:

$$3/5 \square\ P2/P1 \square\square\ 2/1 \quad P_1 \neq 0$$

$$\square\square\square\square\square\square P2/P1 \square\square\ 2 \quad P_1 \neq 0$$

Таким образом найденное решение будет оптимальным, пока отношение цены продукции А к цене продукции В будет находиться в диапазоне от 0,6 до 2.

Задача 2 (Многокритериальная задача)

Используя условие задачи 1, найти план работы при котором достигается:

А) Максимум дохода

Б) Минимум затрат ресурсов (в натуральном выражении)

В) Максимум выпуска продукции А в натуральном выражении

Задача решается методом уступок Величина уступок выбирается студентом.

Решение

Как было показано в задаче 1, максимум выручки V = P$1X_1$ + P2X_2 = 3X_1 + 2X_2 → max достигается в точке B (15, 75).

Минимум затрат ресурсов определяется минимумом целевой функции:

R= (3+4+2)x_1 + (5+1+6)x_2 = 9x_1+12x_2 → min

Поскольку ограничения на минимальный объем продукции не заданы, то минимум затрат ресурсов будет достигаться при полном прекращении выпуска продукции, т.е. когда $x_1 = 0$ и $x_2 = 0$. Это же видно из рассмотрения области OABC на рис. 1. Соответственно минимум функции затрат ресурсов R=0.

В оптимальной по критерию максимума выручки точке B (10,0) целевая функция принимает значение:

V= 3$x1$+2$x2$ =3*10+2*0 =30

Примем величину уступки 90%

90%V=30*0,9 =27

То есть

V= 3$x1$+2$x2$ =27

Нанесем прямую 3$x1$+2$x2$ =27 на график (рис. 2)

Для поиска минимума функции R=9$x1$+12$x2$ построим вектор M{9;12}. Его проекция на ось OX_1 равна 9, на ось OX_2 12.

Поскольку необходимо найти минимум функции R, будем перемещать прямую m, перпендикулярно вектору M, от конца к началу вектора M, т.е. в направлении уменьшения функции R. Перейдя в точку K, прямая m окажется на выходе из области KBP. Точка K – крайняя точка прямой 3$x1$+2$x2$ =27 в области OABC при движении в направлении к началу вектора M, поэтому значение функции R в этой точке будет наименьшим по сравнению с ее значениями в других точках области.

Решив систему уравнений:

□□□ □□□□□□x_1 +5x2 □□□□

□□□ □□□□

□□□ □□□□□□X_1□□X_2□□□□

Найдем x_1 = 8 1/3

x_2 = 1

Таким образом решение многокритериальной задачи при уступке по максимуму выручки 90% - точка К(8 1/3; 1).

Задача 3 (Принятие решений в условиях неопределенности)

Магазин лродает скоропортящуюся продукцию по A рублей за ящик, закупая ее у поставщиков по B рублей за ящик. Непроданная в течение дня продукция реализуется в конце дня по C рублей за ящик. Суточный спрос на продукцию колеблется от 0 до 10 ящиков. Других сведений о спросе нет. Сколько ящиков продукции должен закупать у оптовиков магазин ежедневно в соответствии с принципами максимакса, максимина и минимакса.

Вариант

N	A	в	C
12	50	20	5

Решение

Матрица прибыли (платежная матрица)

		Объем спроса										
		0	1	2	3	4	5	6	7	8	9	10
ок	1	-15	30	30	30	30	30	30	30	30	30	30

	2	-30	15	60	60	60	60	60	60	60	60	60
	3	-45	0	45	90	90	90	90	90	90	90	90
	4	-60	-15	30	75	120	120	120	120	120	120	120
	5	-75	-30	15	60	105	150	150	150	150	150	150
	6	-90	-45	0	45	90	135	180	180	180	180	180
	7	-105	-60	-15	30	75	120	165	210	210	210	210
	8	-120	-75	-30	15	60	105	150	195	240	240	240
	9	-135	-90	-45	0	45	90	135	180	225	270	270
	10	-150	-105	-60	-15	30	75	120	165	210	255	300

Применив критерий Maximax, найдем такой объем закупок, при котором прибыль магазина максимальна при наиболее благоприятном спросе.

Применив критерий Maximax, найдем такой объем закупок, при котором прибыль магазина максимальна при наиболее благоприятном спросе.

		Объем спроса											
		0	1	2	3	4	5	6	7	8	9	10	MAX
Объем закупок	1	-15	30	30	30	30	30	30	30	30	30	30	30
	2	-30	15	60	60	60	60	60	60	60	60	60	60
	3	-45	0	45	90	90	90	90	90	90	90	90	90
	4	-60	-15	30	75	120	120	120	120	120	120	120	120
	5	-75	-30	15	60	105	150	150	150	150	150	150	150
	6	-90	-45	0	45	90	135	180	180	180	180	180	180
	7	-105	-60	-15	30	75	120	165	210	210	210	210	210
	8	-120	-75	-30	15	60	105	150	195	240	240	240	240
	9	-135	-90	-45	0	45	90	135	180	225	270	270	270
	10	-150	-105	-60	-15	30	75	120	165	210	255	300	300

Таким образом, по критерию Maximax оптимально продавать 30 ящиков.

Применим критерий Maximin (Вальда), найдем такой объем закупок, при котором прибыль магазина за неделю максимальна (убыток минимален) при самых неблагоприятных условиях спроса.

Применим критерий Maximin (Вальда), найдем такой объем закупок, при котором прибыль магазина за неделю максимальна (убыток минимален) при самых неблагоприятных условиях спроса.

	Объем спроса											
	0	1	2	3	4	5	6	7	8	9	10	MIN

Объем закупок	1	-15	30	30	30	30	30	30	30	30	30	30	-15
	2	-30	15	60	60	60	60	60	60	60	60	60	-30
	3	-45	0	45	90	90	90	90	90	90	90	90	-45
	4	-60	-15	30	75	120	120	120	120	120	120	120	-60
	5	-75	-30	15	60	105	150	150	150	150	150	150	-75
	6	-90	-45	0	45	90	135	180	180	180	180	180	-90
	7	-105	-60	-15	30	75	120	165	210	210	210	210	-105
	8	-120	-75	-30	15	60	105	150	195	240	240	240	-120
	9	-135	-90	-45	0	45	90	135	180	225	270	270	-135
	10	-150	-105	-60	-15	30	75	120	165	210	255	300	-150

Таким образом, по критерию Maximin (Вальда), оптимально закупать -15 ящиков.

Применив критерий Minimax определим такой объем закупок, при котором риск магазина (упущена выгода) минимален при самых неблагоприятных условиях спроса.

Записав платежную матрицу:

Применив критерий Minimax определим такой объем закупок, при котором риск магазина (упущена выгода) минимален при самых неблагоприятных условиях спроса.												
		Объем спроса										
		0	1	2	3	4	5	6	7	8	9	10
Объем закупок	1	-15	30	30	30	30	30	30	30	30	30	30
	2	-30	15	60	60	60	60	60	60	60	60	60
	3	-45	0	45	90	90	90	90	90	90	90	90

	4	-60	-15	30	75	120	120	120	120	120	120	120
	5	-75	-30	15	60	105	150	150	150	150	150	150
	6	-90	-45	0	45	90	135	180	180	180	180	180
	7	-105	-60	-15	30	75	120	165	210	210	210	210
	8	-120	-75	-30	15	60	105	150	195	240	240	240
	9	-135	-90	-45	0	45	90	135	180	225	270	270
	10	-150	-105	-60	-15	30	75	120	165	210	255	300
MAX		**-15**	**30**	**60**	**90**	**120**	**150**	**180**	**210**	**240**	**270**	**300**

Составим матрицу рисков.

Применив критерий Maximax, найдем такой объем закупок, при котором прибыль магазина максимальна при наиболее благоприятном спросе.												
		Объем спроса										
		0	1	2	3	4	5	6	7	8	9	10	MAX
Объем закупок	1	0	0	30	60	90	120	150	180	210	240	270	270
	2	15	15	0	30	60	90	120	150	180	210	240	240
	3	30	30	15	0	30	60	90	120	150	180	210	210
	4	45	45	30	15	0	30	60	90	120	150	180	180
	5	60	60	45	30	15	0	30	60	90	120	150	150
	6	75	75	0	45	30	15	0	30	60	90	120	120
	7	90	90	75	60	45	30	15	0	210	60	90	210
	8	105	105	90	75	60	105	30	15	0	30	60	105

	9	120	120	105	90	75	60	45	30	15	0	30	120
	10	135	135	120	105	90	75	60	45	30	15	0	135

С точки зрения критерия минимаксного риска Сэвиджа оптимальна стратегия, при которой величина риска минимальна – 30, т.е. оптимальное количество закупаемых ящиков – 13 шт.

Задача 4 (Принятие решений в условиях риска)

Основываясь на условиях задачи 3, определить количество закупаемых магазином для продажи ящиков продукции если известны данные о продажах за последние пятьдесят дней.

Количество проданных ящиков	1	2	3	4	5	6	7	8	9	10	0
Количество дней продаж	2	3	5	5	7	8	7	5	4	2	2

Решение

Рассчитаем вероятности спроса ящиков как доли от общего количества дней продажи.

Количество проданных ящиков	1	2	3	4	5	6	7	8	9	10	0	Итого
Количество дней продаж	2	3	5	5	7	8	7	5	4	2	2	50
Вероятность спроса	0,04	0,06	0,1	0,1	0,14	0,16	0,14	0,1	0,08	0,04	0,04	1

Составим матрицу.

		Вероятность спроса											Средняя прибыль P
		0,04	0,06	0,1	0,1	0,14	0,16	0,14	0,1	0,08	0,04	0,04	
		Объем спроса											
		0	1	2	3	4	5	6	7	8	9	10	
Объем закупок	1	-15	30	30	30	30	30	30	30	30	30	30	28,2
	2	-30	15	60	60	60	60	60	60	60	60	60	53,7
	3	-45	0	45	90	90	90	90	90	90	90	90	74,7
	4	-60	-15	30	75	120	120	120	120	120	120	120	91,2
	5	-75	-30	15	60	105	150	150	150	150	150	150	101,4
	6	-90	-45	0	45	90	135	180	180	180	180	180	104,4
	7	-105	-60	-15	30	75	120	165	210	210	210	210	101,1
	8	-120	-75	-30	15	60	105	150	195	240	240	240	93,3
	9	-135	-90	-45	0	45	90	135	180	225	270	270	81,9
	10	-150	-105	-60	-15	30	75	120	165	210	255	300	68,7

Максимальное значение принимает средняя прибыль для объема закупок 6 ящиков – 104,4.

3.7 Практическое занятие № 7

Тема: «Графический метод решения задач линейного программирования. Задачи линейного программирования»

Решить графически задачу линейного программирования.

$$\begin{cases} 2x_2 - 3x_1 \le 3, \\ 2x_2 + 3x_1 \le 21, \\ x_1 - x_2 \le 2, \\ x_2 + x_1 \ge 4, \\ x_1 \ge 0, \quad x_2 \ge 0, \end{cases}$$

$$f = 3x_1 + x_2 \to min.$$

Решение задачи графическим методом

В первую очередь, найдем область допустимых значений, т.е. точки x1 и x2, которые удовлетворяют системе ограничений. По условию задачи x1 $\ge$ 0, x2 $\ge$ 0, т.е. мы рассматриваем только те точки, которые принадлежат первой четверти. Рассмотрим неравенство 1 системы ограничений.

3x1+2x2$\le$ 3

Преобразуем уравнение следующим образом (разделим на 3):

x1+2/3x2 $\le$ 1

Знак неравенства меньше или равно нуля, следовательно, нас интересуют точки лежащие ниже построенной нами прямой. Рассмотрим неравенство 2 системы ограничений.

3x1+2x2$\le$ 21, что эквивалентно

$$\frac{x_1}{1/3} + \frac{x_2}{1/2} \le 21$$

Преобразуем уравнение следующим образом (разделим на 21):

$$\frac{x_1}{7} + \frac{x_2}{21/2} \le 1$$

Знак неравенства меньше или равно нуля, следовательно, нас интересуют точки лежащие ниже построенной нами прямой.

Рассмотрим неравенство 3 системы ограничений. x1-x2$\le$ 2, что эквивалентно

Преобразуем уравнение следующим образом (разделим на 2):

$$\frac{x_1}{2} - \frac{x_2}{2} \le 1$$

Знак неравенства меньше или равно нуля, следовательно, нас интересуют точки лежащие ниже построенной нами прямой.

Рассмотрим неравенство 4 системы ограничений. x1+x2 $\ge$ 4, что эквивалентно

Преобразуем уравнение следующим образом (разделим на 4):

$$\frac{x_1}{4} - \frac{x_2}{4} \ge 1$$

Знак неравенства больше или равно нуля, следовательно, нас интересуют точки лежащие выше построенной нами прямой.

Построим область допустимых решений, т.е. решим графически систему неравенств. Для этого построим каждую прямую и определим полуплоскости, заданные неравенствами (полуплоскости обозначены штрихом).

Пересечением полуплоскостей будет являться область, координаты точек которого удовлетворяют условию неравенствам системы ограничений задачи.

Обозначим границы области многоугольника решений.

Построим прямую, отвечающую значению функции F = 3x1+x2 = 0. Будем двигать эту прямую параллельным образом. Поскольку нас интересует минимальное решение, поэтому двигаем прямую до первого касания обозначенной области.

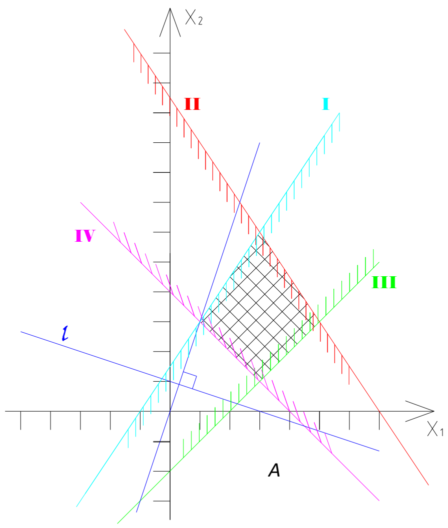


Рисунок 1.1 - Графический метод решения линейного уравнения

Прямая F (x) = const пересекает область в точке A. Так как точка A получена в результате пересечения прямых (1) и (4), то ее координаты удовлетворяют уравнениям этих прямых:

3x1+2x2≤3

x1+x2≥4

Решив систему уравнений, получим: x1 = 1, x2 = 3

Откуда найдем минимальное значение целевой функции:

F (X) = 3*1 + 1*3 = 6

3.8 Практическое занятие № 8

Тема: «Симплексный метод решения задач линейного программирования»

Симплекс метод - метод линейного программирования, который реализует рациональный перебор базисных допустимых решений, в виде конечного итеративного процесса, необходимо улучшающего значение целевой функции на каждом шаге.

Для того чтобы задача имела оптимальное решение необходимо:

– чтобы ограничения задачи были совместимы;

– целевая функция была ограничена сверху при поиске максимального значения и снизу при поиске минимального значения.

Для решения задачи линейного программирования симплекс-методом необходимо привести задачу к канонической форме, для этого все ограничения должны иметь вид строгого равенства:

– ограничения вида «меньше или равно» приводятся к строгому равенству путём добавления к левой части каждого ограничения неотрицательной целой переменной, которая записывается в целевую функцию с коэффициентом равным нулю;

– чтобы привести ограничения вида «больше или равно» к строгому равенству надо из левой части каждого ограничения вычесть неотрицательную целую переменную и добавить искусственную неотрицательную переменную (y1,y2…yn), которая войдёт в целевую функцию с коэффициентом (-M).

Алгоритм:

1) Привести математическую модель к канонической форме, заполнить симплекс таблицу и вычислить значение целевой функции и относительных оценок Δi.

2) Если все относительные оценки не отрицательны и среди базисных переменных нет искусственных, то план оптимален и задача решена. Решение находится в столбце B:

$x_{n+1} = b_1$;

$x_{n+2} = b_2$;

$x_{n+m} = b_m$.

Если искомая переменная не вышла в базис, значит, она равна нулю.

Если все относительные оценки не отрицательны и среди них есть хотя бы одна равная нулю, а в базисе нет искусственных переменных, то задача решена и имеет бесконечное множество решений, одно из которых вышло в базис и находится в столбце B;

Если все $\Delta i \geq 0$, а в базисе остались искусственные переменные, то задача решений не имеет;

Если среди относительных оценок есть отрицательные, то данный план не оптимален и необходим переход к следующему базисному плану.

3) Из всех относительных оценок выбирается наибольшая по модулю, столбец соответствующий ей называется главным. Для определения главной строки надо поэлементно разделить столбец B на главный столбец, наименьшее из получившихся частных соответствует главной строке. Главный элемент симплекс-таблицы находится на пересечении главной строки и главного столбца. Далее необходимо произвести пересчёт симплекс-таблицы.

4) В новой итерации элементы главной строки делятся на главный элемент, а элементы главного столбца делят на «минус главный элемент». Если в главной строке или в главном столбце есть ноль, то элементы, соответствующие этим нулям столбца и строки переписываются в новую таблицу без изменений. Остальные клетки симплекс-таблицы пересчитываются по правилу прямоугольника - из произведения элементов главной диагонали, проходящей через главный элемент, вычитают произведение элементов побочной диагонали и результат делят на главный элемент.

5) Вычисляется значение целевой функции, оно определяется путём вычисления суммы парных произведений столбца C и столбца B. Значение относительных оценок так же вычисляются как сумма парных произведений элементов столбца C на соответствующий столбец.

6) Через конечное число итераций либо будет получено решение задачи линейного программирования, либо будет установлено, что решение неограниченно.

3.9. Практическое занятие №9 (2 часа)

Тема: «Графический метод решения задач линейного программирования Задачи линейного программирования »

Симплекс метод - метод линейного программирования, который реализует рациональный перебор базисных допустимых решений, в виде конечного итеративного процесса, необходимо улучшающего значение целевой функции на каждом шаге.

Для того чтобы задача имела оптимальное решение необходимо:

– чтобы ограничения задачи были совместимы;

– целевая функция была ограничена сверху при поиске максимального значения и снизу при поиске минимального значения.

Для решения задачи линейного программирования симплекс-методом необходимо привести задачу к канонической форме, для этого все ограничения должны иметь вид строгого равенства:

– ограничения вида «меньше или равно» приводятся к строгому равенству путём добавления к левой части каждого ограничения неотрицательной целой переменной, которая записывается в целевую функцию с коэффициентом равным нулю;

– чтобы привести ограничения вида «больше или равно» к строгому равенству надо из левой части каждого ограничения вычесть неотрицательную целую переменную и добавить искусственную неотрицательную переменную (y1,y2…yn), которая войдёт в целевую функцию с коэффициентом (-M).

Алгоритм:

7) Привести математическую модель к канонической форме, заполнить симплекс таблицу и вычислить значение целевой функции и относительных оценок Δi.

8) Если все относительные оценки не отрицательны и среди базисных переменных нет искусственных, то план оптимален и задача решена. Решение находится в столбце B:

$x_{n+1} = b_1$;

$x_{n+2} = b_2$;

$x_{n+m} = b_m$.

Если искомая переменная не вышла в базис, значит, она равна нулю.

Если все относительные оценки не отрицательны и среди них есть хотя бы одна равная нулю, а в базисе нет искусственных переменных, то задача решена и имеет бесконечное множество решений, одно из которых вышло в базис и находится в столбце B;

Если все $\Delta i \geq 0$, а в базисе остались искусственные переменные, то задача решений не имеет;

Если среди относительных оценок есть отрицательные, то данный план не оптимален и необходим переход к следующему базисному плану.

9) Из всех относительных оценок выбирается наибольшая по модулю, столбец соответствующий ей называется главным. Для определения главной строки надо поэлементно разделить столбец B на главный столбец, наименьшее из получившихся частных соответствует главной строке. Главный элемент симплекс-таблицы находится на пересечении главной строки и главного столбца. Далее необходимо произвести пересчёт симплекс-таблицы.

10) В новой итерации элементы главной строки делятся на главный элемент, а элементы главного столбца делят на «минус главный элемент». Если в главной строке или в главном столбце есть ноль, то элементы, соответствующие этим нулям столбца и строки переписываются в новую таблицу без изменений. Остальные клетки симплекс-таблицы пересчитываются по правилу прямоугольника - из произведения элементов главной диагонали, проходящей через главный элемент, вычитают произведение элементов побочной диагонали и результат делят на главный элемент.

11) Вычисляется значение целевой функции, оно определяется путём вычисления суммы парных произведений столбца C и столбца B. Значение относительных оценок так же вычисляются как сумма парных произведений элементов столбца C на соответствующий столбец.

12) Через конечное число итераций либо будет получено решение задачи линейного программирования, либо будет установлено, что решение неограниченно.

4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ

ПО ПРОВЕДЕНИЮ СЕМИНАРСКИХ ЗАНЯТИЙ (не предусмотрены рабочей программой