

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ – МСХА
имени К.А. Тимирязева

А.В.Смиряев, А.В.Исачкин, Л.К.Панкина

МОДЕЛИРОВАНИЕ
В БИОЛОГИИ И СЕЛЬСКОМ ХОЗЯЙСТВЕ

Учебное пособие

Издательство РГАУ-МСХА
Москва
2015 г.

УДК 57.001.57+33.001.57

ББК 28.03яб

C50

Рецензенты: доктор физ.-мат. наук профессор кафедры биофизики биологического факультета МГУ имени М.В. Ломоносова, зав. Сектором информатики и биофизики сложных систем **Г.Ю. Ризниченко**; д.б.н. профессор кафедры генетики, биотехнологии, селекции и семеноводства, руководитель центра молекулярной биотехнологии РГАУ-МСХА имени К.А. Тимирязева **Г.И. Карлов**.

Смиряев А.В., Исачкин А.В., Панкина Л.К.

C50 Моделирование в биологии и сельском хозяйстве: уч. пособие. Издание 3-е исправленное / Смиряев А.В., Исачкин А.В., Панкина Л.К. – М.: Издательство РГАУ-МСХА, 2015. – 153 с.

ISBN 978-5-9675-0824-0

В учебном пособии изложены принципы современного моделирования: основные понятия, классификация моделей и методов моделирования, их возможности и ограничения. Материал иллюстрирован примерами применения моделирования и задачами (большинство со схемами решения) из теории эволюции, экологии, генетики, биотехнологии, селекции, растениеводства, физиологии и защиты растений, медицины, вирусологии, радиологии, демографии, а также из экономики.

Предназначено для подготовки бакалавров по направлениям 110400.62 «Агрономия», 270700.62 «Биотехнология» и 110500.62 «Садоводство».

УДК 57.001.57+33.001.57

ББК 28.03яб

ISBN 978-5-9675-0824-0

© ФГБОУ ВПО РГАУ – МСХА
имени К.А. Тимирязева, 2015
© Издательство РГАУ – МСХА
имени К.А. Тимирязева, 2015

Введение

Моделирование как метод исследования все шире используется в различных областях знаний: от биологии до астрономии, от экономики до медицины и демографии. Причем методы моделирования во многом сходны, хотя и учитывают специфику объекта моделирования.

Так, чрезвычайная сложность биологических систем заставляет с осторожностью относиться к данным, полученным при использовании их моделей. Поэтому анализ результатов моделирования должен сопровождаться тщательным сопоставлением со сведениями об оригинале. Это позволяет не только выявить те звенья причинно-следственной цепи, которые ускользают от исследователя при изучении модели, но и органически включить моделируемые свойства в целостное функционирование живых систем. Возникает вопрос о корректном использовании математических моделей и правильной интерпретации результатов моделирования.

Специфичность биологических систем требует применения адекватного математического аппарата. Однако это вовсе не значит, что необходимо ждать появления новой биологической математики. В биологических исследованиях накоплен обширный опыт использования существующих математических методов и моделей.

Сложность математических моделей с неизбежностью ведет к широкому использованию компьютерной техники как для обработки данных и уточнения параметров моделей, так и для постановки машинного эксперимента, во многих случаях призванного заменить дорогостоящий натуральный эксперимент. Поэтому дальнейшее развитие математического моделирования видится на пути создания новых информационных технологий как инструмента построения содержательных моделей, накопления и хранения информации, полученной в результате исследования этих моделей.

Цель учебного пособия – познакомить студентов различных, прежде всего сельскохозяйственных направлений и специальностей с основными идеями, методами, возможностями и ограничениями современного моделирования в широком диапазоне применения. Материал иллюстрирован упрощенными примерами из теории эволюции, экологии, генетики, биотехнологии, селекции, растениеводства, физиологии и защиты растений, медицины, вирусологии, радиологии, демографии, а также из экономики. Более подробное изложение, как теории моделирования, так и примеров можно найти в литературе, список которой приведен в конце учебного пособия.

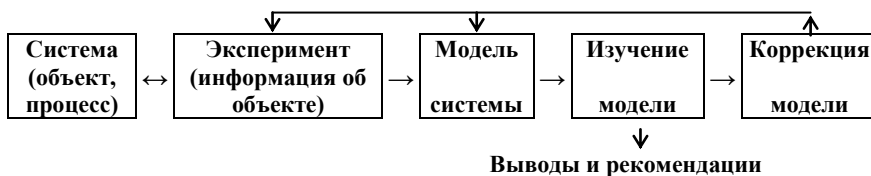
Разделы 1-6 подготовлены проф. А.В.Смиряевым и к.б.н. Л.К.Панкиной, 7 раздел – проф. А.В.Исачкиным.

1. Модели и моделирование

1.1 Задачи моделирования

Моделирование – это процесс построения и изучения модели какого-либо объекта (системы, процесса).

Существуют различные определения модели. Одно из наиболее общих определений следующее: модель – это материальный или мыслительный (абстрактный) объект, который по ходу изучения замещает объект – оригинал (процесс), сохраняя некоторые свойства последнего, важные для конкретного исследования. Последовательность этапов построения модели, взаимосвязь модели и объекта можно представить следующим образом:



Например, вместо того, чтобы сразу строить новый самолёт, сначала строят его модель и помещают его в аэродинамическую трубу. Затем анализируют полученные параметры «полёта», устанавливают связи между параметрами, которые выражают через формулы. После чего вносят определённые коррективы в модель самолёта, проводят дополнительные эксперименты и т.д. По результатам моделирования формулируют окончательные рекомендации проектировщикам самолёта.

Модели используют для того, чтобы:

1. понять, какова внутренняя структура конкретного объекта или (и) структура его взаимодействия со средой;
2. установить наиболее важные связи внутри структуры;
3. установить количественные параметры этих связей;
4. прогнозировать изменения объекта и среды при определенных воздействиях;
5. провести оптимизацию объекта и (или) внешних воздействий на него.

Удачная модель даёт новые знания об объекте, причём относительно дешево и быстро.

1.2. Классификация моделей

Модели можно подразделить на **физические** и **аналоговые**. Например, физические модели: 1) планетарий – физическая модель вселенной, 2) лотки с водой - физическая модель гидроэлектростанции.

Примером аналоговой модели может служить маятник для изучения колебаний основных характеристик электрической цепи переменного тока, поскольку оба процесса описываются одинаковыми уравнениями колебаний.

В рамках аналоговых моделей выделяют **знаковые**, которые включают в себя **математические модели**. В данном случае определение модели звучит так: **математическая модель – это система упрощающих предположений об объекте, допускающих математическую формализацию и применяемых, когда точные закономерности неизвестны или сложны**. В свою очередь математические модели подразделяются на дескриптивные (описательные) и оптимизационные.

Дескриптивные модели служат для описания и прогнозирования характеристик объекта (процесса, системы), например, урожая той или иной сельскохозяйственной культуры от агротехники, почвенно-климатических условий выращивания и других факторов.

Цель **оптимизационных моделей** – найти оптимальное воздействие на объект (процесс). Например, определить оптимальную агротехнику для конкретного сорта сельскохозяйственной культуры

Можно построить несколько моделей одного объекта. Модель называется **адекватной**, если она соответствует данным, полученным в реальных экспериментах с объектом (процессом). Адекватность модели проверяют с помощью методов математической статистики.

Как отмечено выше любая математическая модель работает в рамках определенных предположений, невыполнение которых может привести к ошибочным выводам об объекте. Эти предположения никогда в точности не выполняются, но разные модели обладают различной **робастностью** – устойчивостью к невыполнению предположений. То есть, при не слишком больших отклонениях от предположений выводы и рекомендации, полученные на основе робастных моделей, оказываются верными. Робастными являются, например, модели дисперсионного анализа. В них, кроме прочего, предполагается нормальное (гауссово) распределение изучаемых случайных величин, равенство их дисперсий и пр. Естественно, таких идеальных случаев в природе не бывает, но рекомендации дисперсионного анализа обычно справедливы при некоторых отклонениях от предположений.

Рассмотрим пример построения дескриптивной модели конкретного объекта – популяции рыб, запущенных в озеро, для определения прогноза её численности.

Пусть $x(t)$ – численность рыб в момент времени t ;

$x(0)=x_0$ - численность рыб в начальный момент времени.

Вначале необходимо осуществить **идентификацию** модели – подобрать структуру уравнения (системы уравнений).

Естественно предположить, что в первые годы, когда питания и пространства для каждой особи достаточно, скорость роста численности пропорциональна самой численности x : $dx/dt=kx$, где k – коэффициент пропорциональности. То есть, чем больше численность рыб, тем больше в единицу времени они оставляют потомства (больше скорость роста популяции). Но постепенно с ростом x из-за перенаселения озера возникает ограничение скорости роста численности, которое упрощенно считаем пропорциональным частоте встречаемости рыб: $a(x \cdot x)$. Здесь a – коэффициент пропорциональности.

Итак, можно составить простое дифференциальное уравнение, известное как уравнение Бернулли:

$$dx/dt = kx - ax^2$$

Из курса высшей математики получаем его решение:

$$x(t) = \frac{kx_0 e^{kt}}{k - ax_0(1 - e^{kt})}$$

Далее проводится **настройка** этой модели – подбор числовых значений её параметров (a и k). Их иногда удается получить из литературных данных. Другой путь – провести серию предварительных экспериментов с объектом (например, получить несколько значений x для разных t) и построить экспериментальную кривую зависимости $x(t)$. Далее, используя вышеприведенное решение уравнения Бернулли, методом наименьших квадратов определяют коэффициенты a и k , то есть подбирают их оценки, обеспечивающие минимальные отклонения экспериментальных данных от прогноза по формуле $x(t)$.

Полученную модель можно использовать для прогноза численности рыб (x) через определенные промежутки времени (t).

Для повышения надежности любой модели весьма полезным приемом является ее **верификация**. Общую выборку экспериментальных данных делят на две части. По первой проводят идентификацию и настройку модели. Далее сопоставляют прогнозы, полученные по этой модели, с экспериментальными данными второй части: модель надежна, если результаты статистически близки. Единственное ограничение применения метода верификации – общий

объем выборки экспериментальных данных должен быть достаточным как для построения, так и проверки модели.

Если целью моделирования является не просто описание и прогнозирование процесса, а поиск **оптимальных воздействий** на этот процесс, то в модели из всех параметров, влияющих на изучаемый процесс, выделяют те, на которые человек может воздействовать. Это так называемые **переменные управления** (U). Далее в зависимости от поставленной задачи определяют какие значения и каких выходных параметров системы (процесса) необходимо получить. Желательно, чтобы все выходные параметры были объединены в одну так называемую **целевую функцию** $W(U)$ таким образом, чтобы цель формулировалась просто. Например, добиться максимума $W(U)$ за счет подбора оптимальных значений управляющих воздействий на U . Именно такие модели называются **оптимизационными**, хотя иногда их строят на основе описательных (дескриптивных) моделей.

$$U \rightarrow \boxed{\text{Система}} \rightarrow W(U)$$

Например, можно построить и изучать дескриптивную модель зависимости биомассы популяции рыб (M) от времени (t) с учетом U – отлова рыбы. Но если возникает конкретный вопрос, когда и в каких объемах следует проводить отлов рыбы, чтобы $W(U)$ - суммарный выход отловленной биомассы за 5 лет был максимален, то желательно строить сразу оптимизационную модель.

1.3. Значение моделирования

1. Гипотезы об объекте, выраженные математически, могут служить количественным описанием биологических, сельскохозяйственных и других объектов (процессов) и тем самым способствуют более углубленному их пониманию.

2. Математическая модель часто подсказывает способ представления результатов научных исследований в форме, удобной для использования в практике (графики, гистограммы).

3. Благодаря модели может быть количественно оценена экономическая эффективность внедрения нового объекта или процесса в производство.

4. Модель позволяет выбрать оптимальную стратегию воздействия на объект.

5. Модель дает возможность сократить объем дальнейших экспериментальных работ с объектом.

6. При исследовании сложных объектов модель позволяет объединить разрозненные знания об отдельных частях системы в единое целое (имитационное моделирование).

7. Модель позволяет выбрать наиболее рациональную стратегию и тактику реализации исследовательских программ (теория планирования эксперимента).

8. Математическая модель – это мощное средство обобщения разнородных данных об объекте, позволяющее осуществлять как интерполяцию (восстановление недостающей информации о прошлом), так и экстраполяцию (прогнозирование будущего поведения объекта).

В основе любой математической модели и метода ее анализа лежит математический аппарат. Это могут быть дифференциальные уравнения, формулы теории вероятностей, математической статистики и т.д. Правильный подбор математического аппарата – необходимое условие построения удачной модели.

Контрольные вопросы. 1. Что такое моделирование, общее определение модели, для чего их используют? 2. Приведите классификацию моделей и определение математической модели. 3. В чем разница понятий робастности и адекватности модели? 4. Что такое идентификация, настройка и верификация модели? Как они проводятся? 5. Чем отличаются дескриптивные и оптимизационные модели?

2. Модели динамики биологических систем

2.1. Прогрессия размножения

Приведем несколько приложений дифференциальных уравнений для моделирования биологических систем и процессов.

Ещё Ч. Дарвин обратил внимание на стремление каждого вида к размножению в геометрической прогрессии. То есть, каждая пара организмов дает гораздо больше потомства, чем их выживает до взрослого состояния. Это явление потенциального возрастания численности каждого вида получило название прогрессии размножения. Вся эволюцию можно рассматривать, как ограничение на избыточное размножение.

Рассмотрим модели, с помощью которых можно глубже понять последствия прогрессии размножения и ограничений на неё. Эти несложные математические модели применяются в эволюции, генетике, экологии, биофизике, демографии, медицине и т.д.

Для рассмотрения прогрессии размножения в простейших ситуациях можно не учитывать генетическую структуру популяции, а сконцентрировать все внимание на изменении численности популяции N во времени t . Это первое упрощающее предположение модели.

Основной показатель, характеризующий популяцию определённого вида – **скорость естественного увеличения популяции** (r). Это среднее число потомков, возникающих от одной особи популяции за единицу времени:

$r = b - d$, где b – средняя рождаемость на одну особь за единицу времени; d – средняя смертность в пересчёте на одну особь за единицу времени.

Пример. В начале наблюдения популяция состояла из 800 особей. За год родилось 150 особей (ос./год), а умерло 50. Оценить скорость естественного увеличения популяции (r). Ответ: $r = 150(\text{ос./год}) / 800(\text{ос.}) - 50(\text{ос./год}) / 800(\text{ос.}) = 0,125$ 1/год.

Изменение численности популяции во времени можно выразить через скорость $v = \frac{dN}{dt}$. Сама скорость также как и численность, может меняться во времени. Например, естественным является следующее предположение: чем больше численность популяции в данный момент времени, тем выше скорость увеличения численности (как и в примере с рыбами в начальный период роста популяции).

Основное уравнение, описывающее прогрессивное размножение, когда нет никаких ограничений на N :

$$v = \frac{dN}{dt} = (b - d)N = rN$$

Решение этого уравнения: $N_t = N_0 e^{rt}$, где N_0 – начальная численность популяции; e – основание натурального логарифма.

Графически решение представлено на рис.1. Чем больше r , тем быстрее увеличивается численность популяции.

Еще одно из упрощающих предположений, в рамках которых модель применима – достаточно большое исходное значение N_0 . Только в этом случае можно пренебречь дискретным характером численности.

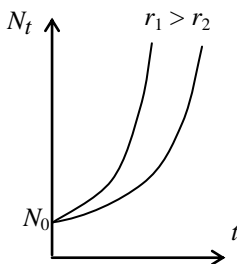


Рис.1. Динамика численности популяции (N) во времени (t) без ограничений численности

Задача. Для первой популяции r - скорость естественного увеличения равна 0,1 1/год, а для второй 0,05 1/год. Начальная численность второй популяции в 2,72 больше начальной численности первой. Определить, через какой промежуток времени численности обеих популяций сравняются.

Схема решения.

$$r_1 = 0,1 \text{ 1/год}; \quad r_2 = 0,05 \text{ 1/год}; \quad N_{02} = 2,72 \cdot N_{01},$$

$$\frac{N_{t2}}{N_{t1}} = 1; \quad \frac{N_{02}e^{r_2 t}}{N_{01}e^{r_1 t}} = 1$$

Ответ: $t = 20$ лет.

Можно привести примеры оценок N_t для разных видов в **предположении отсутствия ограничения на численность**: 1) потомство одной особи бактерии Фишера ($r \approx 100$ 1/час) покроет Землю одним слоем менее чем за 2 суток; 2) за 1300 лет потомство одной пары

слонов образовало бы слой 15 км высотой; 3) потомство одного растения мака при комфортных условиях покрывает всю Землю за 3-4 года.

Подобные потенциальные способности видов к прогрессивному размножению вызывает так называемое «давление жизни», и, как следствие конкуренцию, борьбу за существование между особями, популяциями, видами и т.д. Это фундамент естественного отбора.

Поменяем упрощающее предположение об отсутствии ограничений на N_t – численность популяций. **Ограничения** могут возникать вследствие естественных абиотических и биотических факторов. Наиболее простой вариант (для моделирования) – это стабилизация N_t на некотором максимально допустимом уровне K_{\max} из-за ограниченности ареала обитания популяции.

В модели такое предположение можно отразить, добавив к левой части основного уравнения прогрессии размножения множитель $(1-N/K_{\max})$:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K_{\max}}\right).$$

Получаем уравнение Бернулли. Решение в виде графика зависимости N_t от t представлено на рис.2.

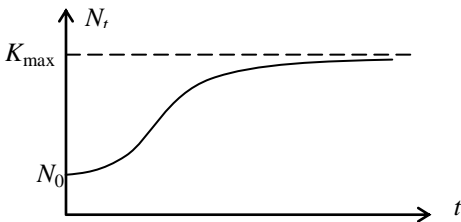


Рис.2. Динамика численности популяции (N) во времени (t) при ограничении на общую численность

При численности N , малой по сравнению с K_{\max} , дополнительный множитель $(1-N/K_{\max})$ близок к 1 и практически не влияет на зависимость N от t . По мере роста численности и приближения N к K_{\max} множитель, а значит и вся правая часть уравнения, приближается к 0. Следовательно, скорость роста популяции также стремится к нулю: кривая роста выходит на плато $N_t = K_{\max}$.

2.2. Моделирование численности взаимодействующих популяций

Предыдущий частный пример - ограничение роста численности абиотическими факторами. Более интересные и важные для эволюции и экологии ситуации возникают при взаимодействии популяций разных видов (перекрывание экологических ниш) или при изменениях внешних условий. В подобных ситуациях по численности N_t возникают так называемые популяционные волны или волны жизни (по С.С Четверикову).

Существует следующая классификация **популяционных волн**:

1. Периодические колебания численности популяции (например, сезонные).
2. Непериодические или периодические колебания численности популяций, вызванные, например, взаимодействием популяций жертвы и хищника.
3. Вспышки численности (когда популяция попадает в благоприятные условия).
4. Резкие сокращения численности популяции (эпифитотии, катастрофы).

Возможны несколько типов взаимодействия двух популяций разных видов: (-, -) – конкуренция, при которых ухудшаются условия жизни для обеих популяций; (+,+) – симбиоз; (+,-) – хищник-жертва или хозяин-патоген и т.д.

Рассмотрим взаимоотношение типа «хищник-жертва». Проследим, как изменится численность популяций хищника и жертвы, если поместить их в ограниченный объём с достаточным количеством пищи для жертв.

Пусть x – численность жертвы; y – численность хищника. Тогда для моделирования можно воспользоваться уравнениями Лотки - Вольтерра:

$$\begin{cases} \frac{dx}{dt} = rx\left(1 - \frac{x}{K_{\max}}\right) - cxy \\ \frac{dy}{dt} = gxy - fy \end{cases}$$

Здесь r – скорость естественного увеличения популяции жертв (без учета влияния хищников);

K_{\max} - предел увеличения численности жертв в ограниченном ареале (число хищников обычно гораздо меньше числа жертв);

cxy – характеризует частоту встречи жертвы и хищника в их ограниченном ареале совместного обитания;

c – коэффициент успеха охоты;
 g – коэффициент рождаемости для хищников (скорость увеличения их численности зависит не только от x , но и y , точнее пропорциональна xy);
 f – коэффициент естественной смертности хищников;

Решения этих уравнений – волновые колебания численности хищника и жертвы. Их форма и периодичность зависят от начальных условий (x_0, y_0) , а также от констант c, f, r, g и K_{\max} .

Возможно несколько вариантов:

1. Выход на стабильный уровень (рис.3) – амплитуда колебаний стремится к нулю. Такая ситуация наблюдается, когда хищникам для $N_t^y = \text{const}$ нужно приблизительно столько жертв, сколько их рождается сверх $N_t^x = \text{const}$.

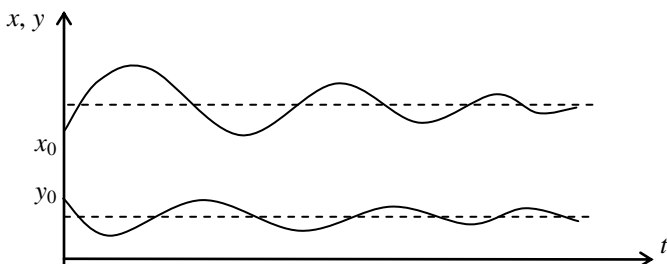


Рис.3. Стабилизация численности жертвы (x) и хищника (y)

2. Интенсивное поедание жертв, а затем и гибель хищника от голода. Волны идут «вразнос» по амплитудам, пока x не станет равным 0 (рис.4).

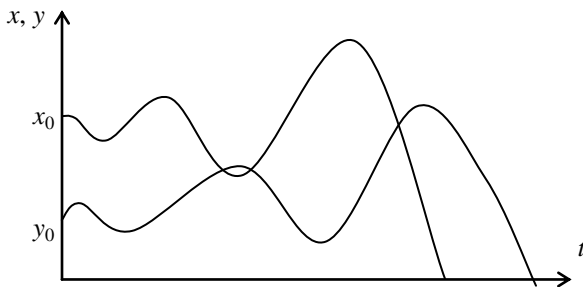


Рис.4. Динамика приближения к гибели жертвы и хищника

3. Стабильные волны с постоянными амплитудами (рис.5).

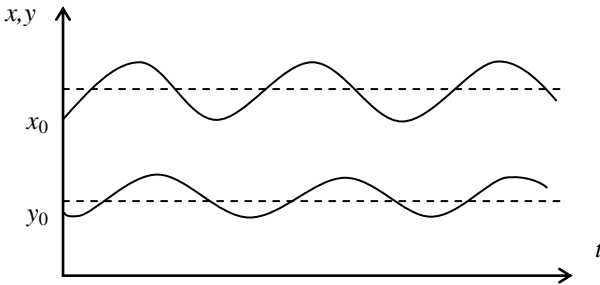


Рис.5. Колебания численности жертвы и хищника с постоянной амплитудой

Эта математическая модель была подтверждена экспериментами на простейших (Смит, 1976), которые послужили физической моделью для изучения обычных экологических и эволюционных ситуаций в природе. Два вида реснитчатых (популяции хищника и жертвы) были помещены в ограниченный объем жидкости (колбу) с добавлением достаточного количества пищи для жертв. Можно было наблюдать разные ситуации:

1. Если в колбе не было хищников, то рост численности популяции жертв происходил до K_{\max} , определяемым объемом жидкости. Форма кривой численности жертв аналогична рис.2.

2. При добавлении в колбу популяции хищника последние активно поедали жертв, увеличивая свою численность. Численность жертв при этом сокращалась, пока они не исчезали полностью, что в свою очередь, приводило к гибели популяции хищника от голода (рис.6).

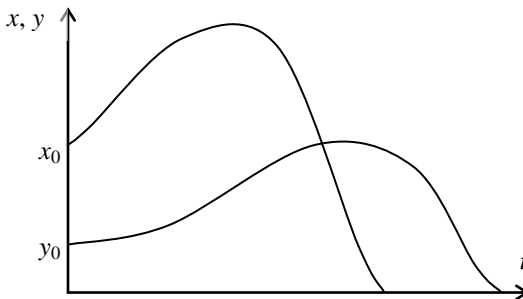


Рис.6. Динамика изменения численности хищника (x) и жертвы (y) для ситуации 2

3. Для того чтобы снизить c – коэффициент успеха охоты, а значит и коэффициент рождаемости для хищника (g), в данном эксперименте в жидкость добавили целлюлозу, которая увеличивала вязкость раствора. В этом случае возникали волны с нарастающей амплитудой до тех пор, пока все жертвы всё же не были съедены ($x=0$), что повлекло за собой гибель хищников (рис.4).

4. Для сокращения скорости естественного увеличения популяции жертв (r) значительно снизили содержание их корма. В этом случае амплитуда увеличения числа жертв стала существенно меньше и, как следствие, не наблюдалось ответного резкого увеличения численности хищника и, в результате, численность жертв резко не снижалась. Возникали стабильные волны по x и y (рис.5).

Аналогичные модели применимы, например, для объяснения и описания колебаний численностей в системе “хозяин – паразит”.

2.3. Модель баланса вещества и энергии

Рассмотрим пример модели, основанный на дифференциальных уравнениях баланса вещества и энергии. Известно, что в природе даже при самых благоприятных условиях рост дерева не превышает некоторого предела. Возникает вопрос, почему деревья любой породы растут сначала быстро, а затем рост замедляется, пока, наконец, совсем не прекращается? Сформулируем гипотезу о причинах этого явления.

Интуитивно ясно, что с ростом кроны, с одной стороны, увеличивается приток энергии благодаря фотосинтезу, а с другой – увеличиваются трудности, связанные, например, с транспортировкой питательных веществ по всему объему дерева и, следовательно, увеличивается расход энергии на подобные нужды. В конце концов, увеличения притока энергии уже не хватает для покрытия расходов, которые увеличиваются быстрее, и дерево перестаёт расти.

На основе этой гипотезы можно сформулировать предположения и построить модель, а затем её исследовать.

Рассмотрим модель И.А. Полетаева. Она основана на следующих упрощающих предположениях:

1. Взрослое растение в процессе роста сохраняет геометрическое подобие. Это значит, что у взрослого растения с ростом не меняются отношения геометрических размеров, например, отношение высоты к диаметру ($h/d = const$).

2. Свободную энергию (или активное вещество) растение получает только путём фотосинтеза.

3. Свободная энергия расходуется на фотосинтез, на подъём раствора из почвы и на построение живой ткани (рост).

4. В среднем за большие отрезки времени растение получает постоянное количество света на единицу поверхности кроны (без учёта суточных и сезонных колебаний) и может поглощать необходимые вещества из достаточного запаса в почве.

Теперь можно составить уравнение баланса.

Пусть x – высота дерева; тогда из 1-го предположения площадь поверхности листьев будет прямо пропорциональна x^2 , а объём растения (например, объём ствола) будет пропорционален величине x^3 . Понятно, что x изменяется со временем: $x = x(t)$. Постараемся выразить все величины, входящие в уравнение баланса, через x .

Сначала найдём выражение для поступающей свободной энергии E . Эта энергия образуется благодаря фотосинтезу. Энергии тем больше, чем больше поверхность зеленой части растения. Таким образом, можно считать, что E пропорциональна x^2 : $E = \lambda x^2$, где λ – коэффициент пропорциональности (он зависит от размеров и формы листьев и от интенсивности фотосинтеза, которые считаем постоянными для конкретной породы).

Других источников энергии, в силу предположения 2 нет, и можно проследить за расходом энергии. Часть энергии, прежде всего, тратится на осуществление самого процесса фотосинтеза. Этот расход также пропорционален x^2 , и его можно представить в виде βx^2 , где β – коэффициент пропорциональности, меньший λ .

Далее энергия расходуется на транспортировку питательного раствора во все части растения. Этот расход будет тем больше, чем больше путей транспортировки, то есть чем больше объём растения. Кроме того, этот расход связан с преодолением силы тяжести и, следовательно, будет тем больше, чем на большую высоту приходится поднимать питательные вещества. Таким образом, этот расход пропорционален как объёму x^3 , так и высоте x , и можно считать, что он пропорционален их произведению, то есть $\gamma x^3 x$.

Наконец, энергия расходуется на увеличение массы растения, то есть на рост. Этот расход пропорционален скорости роста, то есть производной по времени от массы ($m = \rho x^3$, где ρ – средняя плотность растения, x^3 – объём). Таким образом, последний расход может быть выражен как $\delta \frac{d}{dt}(\rho x^3)$, где δ – коэффициент пропорциональности.

В силу закона сохранения энергии и с учётом высказанных предположений расход энергии должен быть равен её притоку, и получаем уравнение баланса:

$$\lambda x^2 = \beta x^2 + \gamma x^4 + \delta \frac{d}{dt}(\rho x^3) \quad \text{или}$$

$$\lambda x^2 = \beta x^2 + \gamma x^4 + \delta \rho 3x^2 \frac{dx}{dt} \quad (1)$$

Это соотношение представляет собой дифференциальное уравнение относительно $x(t)$.

Разделив уравнение (1) на выражение $3\delta\rho x^2$, которое не может быть равным нулю, и обозначив $a = \frac{\lambda - \beta}{3\delta\rho} > 0$; $b = \frac{\gamma}{3\delta\rho} > 0$

получим
$$\frac{dx}{dt} = a - bx^2 \quad (2).$$

Так как дерево растёт, производная dx/dt положительна.

Это значит, что $a - bx^2 > 0$, и следовательно, $x^2 < a/b$. Поэтому, интегрируя выражение (2) при начальной высоте $x(0) \approx 0$, имеем

$$\ln \frac{\sqrt{\frac{a}{b}} + x}{\sqrt{\frac{a}{b}} - x} = 2\sqrt{ab} \times t,$$

откуда, наконец, получаем решение дифференциального уравнения (1):

$$x(t) = \sqrt{\frac{a}{b}} \frac{1 - e^{-2\sqrt{ab} \times t}}{1 + e^{-2\sqrt{ab} \times t}}$$

Эта формула даёт кривую роста дерева от времени (рис. 7).

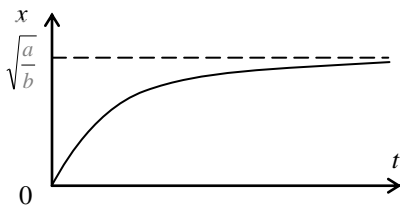


Рис. 7. Динамика роста дерева (x) во времени (t)

Если известны a и b (они зависят, в частности, от породы и сорта дерева), то по этой формуле можно определить средний рост дерева данной породы в зависимости от возраста. Проверка модели в реальных экспериментальных ситуациях подтвердила её адекватность. Следовательно, предположения 1...4, лежащие в ее основе, не противоречат реальности: для объяснения остановки роста дерева достаточно исходной гипотезы.

Задача. Максимальная высота деревьев в лесу 50 м. 40-летние деревья срубают и используют как сырьё для изготовления целлюлозы. Их средняя высота 15 м. Определить коэффициенты a и b (настроить модель).

Схема решения.

При увеличении возраста (t) высота $x(t)$ приближается к $\sqrt{\frac{a}{b}}$

(см. $x(t)$ – решение дифференциального уравнения).

$$\begin{cases} \sqrt{\frac{a}{b}} = 50 \\ \sqrt{\frac{a}{b}} \frac{1 - e^{-2\sqrt{ab}40}}{1 + e^{-2\sqrt{ab}40}} = 15 \end{cases}$$

Решив эту систему уравнений можно найти значения коэффициентов a и b , после чего модель пригодна для получения прогноза средней высоты деревьев любого возраста в лесу.

2.4. Биологический метод борьбы с нежелательным видом

Речь идёт о теоретическом обосновании метода Кюрасао. Сущность этого метода заключается в том, что в популяцию, которую хотят подавить (например, в популяцию сельскохозяйственных вредителей), регулярно вводят стерильных **транс-самцов**. Таких самцов с большим числом транслокаций можно получить, например, подвергнув облучению нормальных самцов. Из-за серьезных нарушений в мейозе у них формируются несбалансированные гаметы. Транс-самцы не оставляют нормального потомства, то есть не участвуют в процессе естественного воспроизводства. В то же время, они вполне жизнеспособны и, наряду с нормальными, участвуют во внутривидовой борьбе, в том числе за самок, снижая тем самым скорость естественного увеличения популяции. При достаточной интенсивности введения в популяцию транс-самцов плотность нормальных самцов в поколениях приближается к нулю.

Рассмотрим модель, предложенную А.Д.Базыкиным.

Пусть $x(t)$ – плотность нормальных самцов на поле (тыс/га).

n^* – постоянная скорость, с которой стерильные самцы вводятся в популяцию (то есть число стерильных особей, вводимых в единицу времени на единицу площади поля);

$y(t)$ – плотность стерильных самцов.

Необходимо определить **минимальную** скорость n^* для постепенного снижения до нуля численности нормальных самцов, то есть достаточную, чтобы $x(t) \rightarrow 0$.

Можно составить уравнения изменения численности нормальных и стерильных самцов,

$$\begin{cases} \frac{dx}{dt} = rx - \delta x^2 - \delta xy \\ \frac{dy}{dt} = n^* - \delta y^2 - \delta xy \end{cases}$$

где $r = b - d$ – это постоянная скорость естественного увеличения плотности нормальных самцов;

δx^2 – снижение скорости роста плотности из-за конкуренции между нормальными самцами;

δy^2 – снижение скорости роста плотности из-за конкуренции между стерильными самцами;

δxy – снижение скорости роста плотности из-за конкуренции между нормальными и стерильными самцами;

Решение этой системы уравнений показало, что $x(t) \rightarrow 0$, если $n^* \geq r^2/\delta$.

Задача. Сколько стерильных самцов необходимо вводить в популяцию насекомых за единицу времени на единицу площади, чтобы $x(t) \rightarrow 0$, если $r = 1$ 1/час, а $\delta = 0,01$ 1/час?

Для решения достаточно подставить эти значения в предыдущую формулу: $n \geq 1/0,01 = 100$ особей в час, т.е. не менее 2400 особей в сутки на единицу площади.

2.5 Модель эпидемии

За многие тысячелетия существования человечества огромное число людей погибло от различных эпидемий. Для того чтобы иметь возможность бороться с эпидемиями, то есть своевременно применять те или иные медицинские мероприятия (карантины, вакцинации и т.д.), необходимо уметь сравнивать эффективность этих мероприятий. Сравнить же их можно лишь в том случае, если есть возможность предсказать, как при том или ином мероприятии будет меняться ход эпидемии, прежде всего число больных. Отсюда возникает необходимость в построении моделей, которые могли бы служить целям прогноза.

Сначала рассмотрим модель «естественного» хода эпидемии (без медицинского вмешательства). Понятно, что модель эпидемии может включать в себя влияние факторов самых различных уровней.

Так, можно было бы учесть законы, управляющие деятельностью бактериальных клеток, степень восприимчивости к инфекции отдельных людей, вероятности встречи носителей инфекции с ещё здоровыми людьми и многие другие факторы. Так как нашей целью является лишь создание иллюстративной модели, то мы абстрагируемся от многих факторов.

Пусть имеется N здоровых людей, и в момент времени $t = 0$ в эту группу попадает один заболевший человек (источник инфекции). Предположим, что никакого удаления заболевших из группы не происходит (нет ни выздоровления, ни гибели, ни изоляции). Будем считать также, что человек становится источником инфекции сразу же после того, как он сам заразится.

Обозначим число заболевших к моменту времени t через $x(t)$, а число здоровых – через $y(t)$ (очевидно, что $x(t)+y(t)=N+1$ в любой момент времени).

При $t = 0$ выполняется условие $x(0) = 1$.

Рассмотрим интервал времени $t+dt$, где dt – малый промежуток времени. Необходимо определить, сколько новых больных (dx) появится за этот промежуток времени. Можно предположить, что их число будет пропорционально величине dt , а также числу встреч здоровых и заболевших людей, то есть произведению величин $x \cdot y$:

$dx = \alpha \cdot x \cdot y \cdot dt$, где α – коэффициент пропорциональности (вероятность передачи инфекции при контакте здорового с больным).

$$\text{Проводим замену } y = N + 1 - x \rightarrow \frac{dx}{dt} = \alpha x(N + 1 - x)$$

Решение этого уравнения:

$$x(t) = \frac{N + 1}{Ne^{-\alpha(N+1)t} + 1}$$

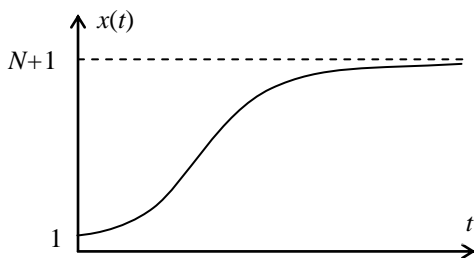


Рис. 8. Динамика числа больных (x) во времени (t)

Полученный прогноз – форма зависимости числа больных в группе от времени представлен на рис.8.

Задача. Оценить количество больных через 6 суток и сколько людей заболит за 6-й день, если $\alpha = 0,001$, а $N + 1 = 1101$ чел.?

Для получения ответов следует использовать решение предыдущего дифференциального уравнения.

Можно усложнить модель, предположив, например, что в начальный момент времени $t=0$ болен не 1 человек, а несколько [$x(0)=b$]. Кроме того, предположим, что через небольшой промежуток времени больной выздоравливает и получает иммунитет. Тогда общее число здоровых к моменту t состоит из y – числа еще не переболевших и $z(t)$ – числа переболевших и выздоровевших. $z(0) = 0$, $y(0) = N$.

$$x + y + z = N + b \quad \rightarrow \quad z = N + b - x - y.$$

$$\begin{cases} \frac{dx}{dt} = \alpha xy - \gamma x \\ \frac{dy}{dt} = -\alpha xy \end{cases}$$

где γx – число выздоровевших.

Прогноз числа больных может иметь разную форму, в частности, представленную на рис. 9. Конкретный вид кривой зависит от N , b , α , γ .

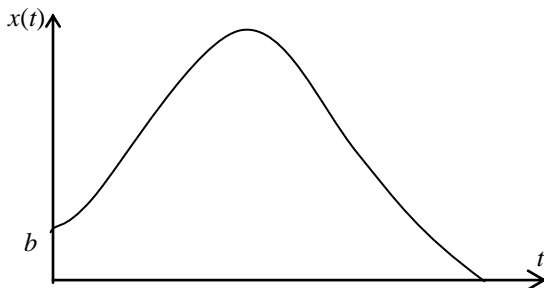


Рис.9. Динамика численности больных во времени с учетом иммунитета

В модели можно учесть смертность от болезни, передачу болезни через переносчика (грызуны), вакцинацию, снижающую α и т.д. Аналогичные модели используются в ветеринарии.

2.6. Модели динамики возрастных групп

Рассмотрим ситуацию, когда вес каждой особи популяции меняется в течение жизни и необходимо сделать прогноз не численности, а биомассы всей популяции или ее «молодой» части через определенное время (t).

Пусть τ - возраст особи; $N(t, \tau)$ – численность всех особей популяции, имеющих в момент времени t возраст τ , $P(\tau)$ – средний вес особи возраста τ .

Тогда биомасса всех особей возраста τ равна $N(t, \tau) P(\tau)$.

Обозначим через $M(t, \varphi)$ интересующую нас биомассу всех особей популяции, имеющих в момент времени t возраст не более φ .

$$\text{Тогда } M(t, \varphi) = \int_0^{\varphi} N(t, \tau) P(\tau) d\tau$$

Зависимость P от τ иногда известна, например, из научной литературы. Гораздо труднее определить численность $N(t, \tau)$. Она зависит от многих факторов как внешних (температура, влажность, питание и т.д.), так и от видовых особенностей (плодовитость, жизнеспособность и т.д.).

Иногда для моделирования численности $N(t, \tau)$ удобно отказаться от непрерывного времени и перейти от дифференциальных уравнений к **дискретным моделям**, прогнозирующим процесс по «шагам», то есть дискретные моменты. Рассмотрим дискретную «шаговую» модель динамики возрастной структуры популяции в зависимости от времени. Эта модель широко используется в научных исследованиях по экологии, сельскому хозяйству, демографии и т.д.

Популяцию условно разбивают на n возрастных групп (рис. 10). В начальный момент времени t_0 известна численность особей каждой (i -й) возрастной группы, которая обозначается $x_i(t_0)$, $i = 1 \dots n$.

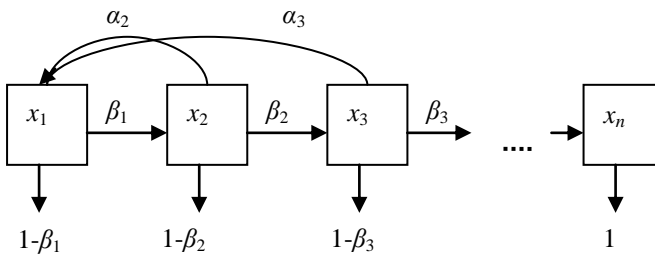


Рис.10. Схема формирования численности возрастных групп

Из всех возрастных групп выделяют те, которые производят потомство. Пусть их номера будут $k, k+1, k+2, \dots, k+p$. Предположим, что следующий момент времени t_1 выбран так, что за промежуток от t_0 до t_1 (1 «шаг»):

- особи i -й возрастной группы переходят в группу $i+1$;
- от групп $k, k+1, k+2, \dots, k+p$ появляется потомство;
- часть особей каждой группы погибает.

Например, первая группа – особи от 0 до 1 года, вторая – от 1 до 2-х лет и т.д. «Шаг» прогноза численностей этих групп также 1 год. Тогда на рис 10:

α_i - коэффициент рождаемости в i -й группе: число потомков i -й возрастной группы, родившихся за один «шаг», в пересчете на одну родительскую особь этой группы;

β_i - коэффициент выживаемости в i -й группе – доля выживших за «шаг» (смертность будет определяться, как $1-\beta_i$). $0 < \beta_i < 1$ для всех групп кроме последней (n -й), где $\beta_n = 0$ (все погибают).

Тогда $x_1(t_1)$ – численность особей первой возрастной группы в момент времени t_1 равна численности особей, родившихся в промежутке от t_0 до t_1 , т.е. равна сумме потомств от всех возрастных групп, производящих потомство за один «шаг». Понятно, что численность потомства от отдельной (i -й) группы будет тем больше, чем больше в ней родителей, то есть чем больше $x_i(t_0)$. Точнее, потомство от i -й группы за один временной «шаг» будет равно $\alpha_i x_i(t_0)$, а все потомство, появившееся в промежутке от t_0 до t_1 будет равно:

$$x_1(t_1) = \sum_{i=k}^{k+p} \alpha_i x_i(t_0)$$

Теперь определим $x_2(t_1)$ – численность второй возрастной группы в момент времени t_1 . За время от t_0 до t_1 особи, находящиеся в момент t_0 в первой возрастной группе перейдут во вторую. За это время часть из них погибнет. Поэтому численность второй возрастной группы в момент времени t_1 : $x_2(t_1) = \beta_1 x_1(t_0)$.

Аналогично рассчитывается и численность каждой (i -й) группы через 1 «шаг»: $x_i(t_1) = \beta_{i-1} x_{i-1}(t_0)$, где $0 < \beta_{i-1} < 1$, $i=2 \dots n$.

Построив прогноз на 1 «шаг» вперед, принимают новые численности $x_i(t_1)$ за исходные и повторяют процедуру, получая прогнозы $x_i(t_2)$ – на 2 «шага» вперед и т.д.. Эти модели можно использовать не только для прогноза на несколько «шагов» вперед, но и для подбора схемы «сбора биомассы» по «шагам» и возрастным группам, обеспечивающей, например, максимальную прибыль за определенный промежуток времени (за несколько «шагов»).

Задача. На птицеферме всех птиц можно отнести к четырем возрастным группам с «шагом» в 1 год. Начальная численность первой возрастной группы – 3000 особей, второй – 1000 особей, третьей и четвертой – по 500 особей. Коэффициент естественной выживаемости для первой и второй возрастной группы – 0,9, для третьей – 0,8. Давать потомство могут только особи второй и третьей возрастных групп. Коэффициент рождаемости для второй возрастной группы, т.е. среднее число потомков от одной особи за год равно 3, а для третьей – 2.

Рассмотрим два варианта ежегодной продажи:

1. Продаём всю четвёртую возрастную группу. Цена 1 кг особей этой возрастной группы 30 руб, а средний вес 4 кг.

2. Продаём всю четвёртую возрастную группу, а также 30% из второй возрастной группы и 40% из третьей возрастной группы. Продажа осуществляется в начале года, до того как особи 2 и 3-ей возрастной группы дадут потомство.

Средний вес особи второй возрастной группы 2 кг при цене 40 руб. за 1 кг. Цена 1 кг особей 3-ей возрастной группы 35 руб., а средний вес 3 кг. Определить прибыли от первого и второго варианта продажи и построить прогноз численности всех возрастных групп на 1 шаг вперёд (через год).

Схема решения.

Прибыль от 1-го варианта продажи составит:

$$П_1 = 500 \text{ шт.} \times 30 \text{ руб.} \times 4 \text{ кг} = 60000 \text{ руб.}$$

Построим прогноз численности на 1 «шаг» вперёд.

С учётом рождаемости численность особей 1-ой возрастной группы через 1 год составит:

$$x_1(t_1) = 3 \cdot 1000 + 2 \cdot 500 = 4000 \text{ шт.}$$

Далее определяем численности 2-ой, 3-ей и 4-ой возрастных групп через год:

$$x_2(t_1) = 3000 \times 0,9 = 2700 \text{ шт.,}$$

$$x_3(t_1) = 1000 \times 0,9 = 900 \text{ шт.,}$$

$$x_4(t_1) = 500 \times 0,8 = 400 \text{ шт.}$$

Теперь определим прибыль от второго варианта продажи.

Число особей, продаваемых из второй возрастной группы при $x_2(t_0) = 1000$ составляет: $1000 \times 0,3 = 300$ шт. Значит во второй возрастной группе остаётся $1000 - 300 = 700$ особей.

Аналогично с 3-ей возрастной группой при $x_3(t_0) = 500$. Продаём: $500 \times 0,4 = 200$ шт. Остаётся $500 - 200 = 300$ особей.

Из 4-ой возрастной группы продаём всех особей, то есть $x_4(t_0) = 500$ шт.

Теперь можно определить прибыль от второго варианта продажи. Она составит:

$$П_2 = 500 \text{ шт.} \times 30 \text{ руб.} \times 4 \text{ кг} + 300 \text{ шт.} \times 2 \text{ кг} \times 40 \text{ руб.} + 200 \text{ шт.} \times 35 \text{ руб.} \times 3 \text{ кг} = 60000 + 24000 + 21000 = 105000 \text{ руб.}$$

Определим численность всех возрастных групп на 1 «шаг» вперёд.

$$x_1(t_1) = 700 \text{ шт.} \times 3 + 300 \text{ шт.} \times 2 = 2700 \text{ шт.}$$

$$x_2(t_1) = 3000 \text{ шт.} \times 0,9 = 2700 \text{ шт.}$$

$$x_3(t_1) = 700 \text{ шт.} \times 0,9 = 630 \text{ шт.}$$

$$x_4(t_1) = 300 \text{ шт.} \times 0,8 = 240 \text{ шт.}$$

Хотя прибыль во втором варианте больше, но численности групп, кроме 2-й, снизились по сравнению с первым вариантом. Это приведёт к снижению прибыли в следующие годы, что можно прогнозировать с помощью той же модели.

Подобные модели используют также в демографии для прогноза ожидаемой численности разных возрастных групп людей на несколько «шагов» (на 5, 10, 15 и т.д. лет) вперёд.

Контрольные вопросы. 1. Приведите уравнение (модель) для описания прогрессии размножения, когда нет никаких ограничений на N . Как изменится эта модель, если ввести ограничение – предельную численность популяции K_{\max} ? 2. Поясните понятие популяционных волн и их классификацию, от чего зависит форма волн численности? 3. Из каких частей состоят уравнения - модель для описания изменений численности популяций хищника и жертвы в их ограниченном ареале совместного обитания? 4. В чем состоит общая гипотеза, объясняющая причину остановки роста дерева, и какие упрощающие предположения используются для построения модели роста? 5. Какова генетическая основа биологического метода борьбы с нежелательным видом? Составьте модель для описания изменений численностей нормальных и стерильных самцов. 6. Приведите модель естественного хода эпидемии при $x(0)=1$, как изменится эта модель, если в момент времени t болен не один человек, а несколько и через небольшой промежуток времени больные выздоравливают и приобретают иммунитет? 7. В чём сложность построения модели для определения биомассы определённых возрастных групп? 8. Сформулируйте демографическую задачу, которая может быть решена с использованием дискретной «шаговой» модели динамики возрастной структуры популяции от времени.

3. Моделирование кинетики метаболизма в биотехнологии

Цель биотехнологии – получение ценных продуктов с использованием биохимической деятельности живых организмов (возможно, модифицированных), изолированных клеток или их компонентов. С точки зрения биотехнологии **метаболизм** – это совокупность химических реакций, обеспечивающих образование и распад биохимических продуктов в процессе развития, размножения и гибели клеток.

В качестве основных видов биохимической деятельности микробиообъектов, используемых в биотехнологии, как правило, выступают следующие:

1. Рост клеточной массы биореагентов, которые и представляют собой продукт. К данному классу процессов относится получение пекарских дрожжей в пищевой промышленности, кормовых дрожжей в сельском хозяйстве, вакцин в медицине.

2. Образование (биосинтез) в процессе роста и развития клеток ценных биохимических продуктов. Некоторые из них выделяются в среду (внеклеточные продукты), некоторые накапливаются в биомассе (внутриклеточные продукты). В этих случаях производство существует ради получения таких продуктов, а не самой биомассы, которая часто является балластом. Пример - получение полисахаридов - внеклеточных продуктов бактерий *Bacillus mucilaginosus* и *Rhizobium leguminosarum*. Эти препараты обладают высоким иммуномодулирующим эффектом.

3. Биотрансформация – процесс, в результате которого под воздействием биологической деятельности микроорганизмов или культуры клеток, возможно, с участием ферментов (биологических катализаторов), происходит изменение химического состава исходного химического вещества. Примерами процесса биотрансформации является превращение глюкозы во фруктозу под воздействием фермента глюкоизомеразы или глицерина в диоксиацетон под воздействием глюконобактерий.

4. Потребление микроорганизмами из жидких сред различных веществ, например, загрязнителей. В этих процессах биомасса микроорганизмов является промежуточным агентом. Такие процессы применяют при биохимической очистке сточных вод.

5. Выщелачивание с помощью микроорганизмов, т.е. перевод в растворенное состояние некоторых веществ, находящихся в твердых телах. Примером данных процессов является микробиологическое выщелачивание металлов из руд в добывающей и металлургической промышленности.

6. Использование биохимической деятельности микроорганизмов с целью образования газов, в частности, для создания

пористых материалов. Подобные технологии используют, например, в пищевой промышленности при производстве хлеба, шипучих напитков (пиво, шампанское) и пр.

3.1. Кинетические характеристики процесса биосинтеза.

В естественных условиях метаболизм настроен так, чтобы производить лишь необходимое количество метаболитов. В биотехнологическом производстве, контролируемом человеком и нацеленном на получение прибыли, такая ситуация неприемлема. Для производственных целей необходимо изучить **кинетику** биохимических реакций т.е. закономерности их протекания во времени. Ввиду сложности названных процессов для их описания и прогнозирования обычно применяют дескриптивные математические модели, основанные на дифференциальных уравнениях.

Кинетические характеристики процесса – количественные показатели протекания биохимических превращений. Основные показатели процесса характеризуют биомассу, субстрат и продукт. Соответственно состояние процесса в простейших случаях определяется следующими показателями:

- плотность (или биомасса) микроорганизмов или культуры клеток в биореакторе: далее она обозначается x ; размерность этого показателя – г/л, г/см³ (или г);
- концентрация питательной среды – субстрата (или его основного компонента) – s , г/л.
- концентрация основного продукта метаболизма – p , г/л; в его качестве могут, выступать не только внеклеточный продукт биосинтеза микроорганизмов, но сама их биомасса, ее компонент (внутриклеточный продукт) и т.д.

В соответствии с этим упрощенным перечнем показателей разработан ряд кинетических моделей, представленных на рис 11.



Рисунок 11. Взаимосвязанные модели биотехнологического процесса

3.2. Моделирование кинетики биомассы

В основе биотехнологического процесса – развитие популяции микроорганизмов во времени. В предыдущей главе отмечалось, что при отсутствии ограничений на рост популяции микроорганизмов, его динамику можно описать простейшим дифференциальным уравнением со структурой:

$$\frac{dx}{dt} = rx,$$

где x – плотность (или биомасса) популяции, r – скорость ее увеличения, t – время.

Решение этого уравнения: $x = x_0 e^{rt}$, где x_0 – начальная плотность популяции; e – основание натурального логарифма.

Если параметр r больше нуля и постоянен во времени, то рост численности (или биомассы, плотности популяции) не ограничен. Но в биореакторе величина r не постоянна, она зависит от меняющихся, отчасти контролируемых условий.

Поясним, что при $r > 0$ плотность популяции (x) растет тем быстрее, чем больше r , при $r = 0$ плотность постоянна, а при $r < 0$ – снижается (рис. 12).

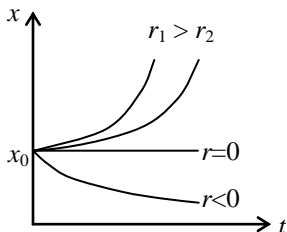


Рис. 12. Зависимость x от t при разных значениях r . Пояснения в тексте

В моделях биотехнологических процессов вместо параметра r , который обычно считается постоянным, часто используют обозначение μ , называемое **удельной скоростью роста** плотности (биомассы) популяции микроорганизмов. При моделировании кинетики биомассы предполагается, что на рост влияют, прежде всего, изменения μ . Итак основное уравнение роста популяции запишем в виде:

$$\frac{dx}{dt} = \mu x \quad (3)$$

Простейший пример – абиотическое ограничение на прогрессию размножения (раздел 2.1), вызванное конечным объемом биореактора. В этом случае можно принять: $\mu=r(1-x/K_{max})$, где K_{max} – максимально допустимая плотность популяции, определяемая этим объемом. Параметр μ в процессе роста популяции меняется от $\mu_0=r(1-x_0/K_{max})$ до $\mu=0$. Решение уравнения (3) при таком характере изменения μ – это S-образная кривая роста x , представленная на рис. 2 раздела 2.1.

Обсудим различные причины изменения параметра μ и их влияние на кинетику плотности популяции.

Влияние концентрации питательной среды – субстрата

В зависимости от конкретного биотехнологического процесса, штамма микроорганизмов и субстрата зависимость μ от s – концентрации конкретного субстрата может иметь самый различный модельный характер. Например, в соответствии с простейшей **моделью Кобозева** удельная скорость роста μ прямо пропорциональна концентрации субстрата, доступного популяции микроорганизмов. Т.е. $\mu=Ks$ (рис. 13), где K – константа, характерная для конкретных условий биотехнологического процесса.

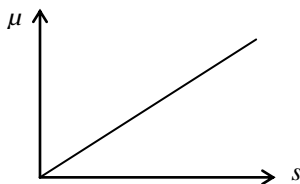


Рис. 13. Зависимость $\mu(s)$ по модели Кобозева.

Следовательно, увеличивая концентрацию (s) субстрата, можно добиться резкого ускорения роста биомассы (x) или наоборот – при снижении s , что следует из уравнения:

$$\frac{dx}{dt} = \mu x = Ksx. \quad (4)$$

В другой **модели Блэкмана** при небольших концентрациях s модельная зависимость та же, что в модели Кобозева ($\mu=Ks$), но при s , больших определенного порога s^* , удельная скорость роста μ уже не меняется:

$$\mu = \begin{cases} Ks & \text{при } s < s^* \\ Ks^* & \text{при } s \geq s^* \end{cases}$$

Для конкретного процесса значение порога подбирается эмпирически – на основе опытных данных. Графически модель Блэкмана представлена на рис. 14.

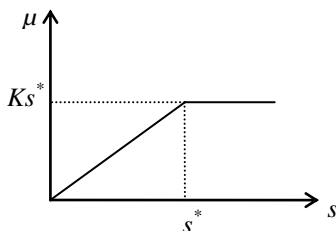


Рис. 14. Зависимость $\mu(s)$ по модели Блэкмана.

Чаще применяются более сложные модели зависимостей $\mu(s)$:

- модель **Моно**

$$\mu = \frac{\mu_m s}{K_s + s}, \quad (5)$$

в которой предполагается, что при увеличении s происходит постепенное приближение удельной скорости роста μ к порогу «насыщения» μ_m (рис. 15 а);

- модель **Мозера**

$$\mu = \frac{\mu_m s^k}{K_s + s^k},$$

предполагает S-образную зависимость $\mu(s)$ (рис. 15 б);

- модель **Андрюса**

$$\mu = \frac{\mu_m s}{K_s + s + s^2 / k}, \quad (6)$$

учитывающая возможность ингибирующего влияния на популяцию завышенных концентраций субстрата (рис. 15 в) и другие модели, которые описаны, например, в работе Д.С. Дворецкого с соавторами (2005г.).

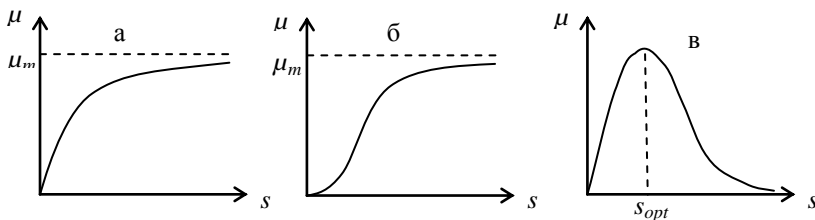


Рис. 15. Зависимость $\mu(s)$ по моделям Моно (а), Мозера (б) и Андрюса (в).

В реальности на удельную скорость роста популяции часто влияет не один, а несколько субстратов. Используют различные многосубстратные модельные зависимости. Чаще всего приходится учитывать влияние двух субстратов (например, углеродного и азотного, углеродного и кислородного). Подбирают одного из четырех основных типов уравнений для параметра μ :

- **Мультипликативные модели**, где общая функция влияния на μ есть **произведение** однофакторных влияний субстратов, причем каждый фактор автономен и может иметь свою собственную зависимость f :

$$\mu = f_1(s_1) f_2(s_2).$$

Например, влияние одного субстрата описывается зависимостью по Моно (5), а второго – с ингибированием по Андрию (6):

$$\mu = \frac{\mu_m s_1}{K_1 + s_1} \frac{\mu_m s_2}{K_2 + s_2 + s_2^2 / k}.$$

- **Аддитивные модели**, где многофакторная функция является **суммой** однофакторных. Такие зависимости встречаются довольно редко, чаще для двух субстратов одного назначения (например, два углеродных субстрата: глюкоза и лактоза, глюкоза и крахмал и т.д.):

$$\mu = \frac{\mu_1 s_1}{K_1 + s_1} + \frac{\mu_2 s_2}{K_2 + s_2}.$$

- **Альтернативные модели**, где зависимость μ от нескольких (k) субстратов подчиняется принципу кинетического минимума:

$$\mu = \min_{i=1}^k \mu_i(s_i).$$

Это уравнение предполагает, что для каждого субстрата существует своя зависимость $\mu_i(s_i)$. Именно она действует, когда лимитирующим фактором является этот субстрат. То есть микроорганизмы растут со скоростью, наименьшей из k возможных зависимостей $\mu_i(s_i)$.

- **Модели с неразделяющимися переменными**. Три предыдущих типа моделей, несмотря на кажущееся различие, сходны в одном. Они формируются из однофакторных зависимостей. Однако возможны и более сложные случаи, когда многофакторную зависимость трудно разбить на однофакторные. Например, существует уравнение «конкурентного торможения» вторым субстратом:

$$\mu = \frac{\mu_m s_1}{K_s + s_1 + s_2 / k}.$$

Особый интерес представляет **модель Контуа**, учитывающая действие на удельную скорость роста популяции одновременно стимулирующего влияния концентрации субстрата и ингибирования роста самой биомассой (x) микроорганизмов (сравните с уравнением (5) – моделью Моно):

$$\mu = \frac{\mu_m s}{Kx + s}.$$

Подбор (идентификация) той или иной модели для описания кинетики конкретного биотехнологического процесса, а также ее настройка (подбор значений констант K , s^* , μ_m , k и пр.) производится на основе литературных сведений и собственных экспериментальных данных. Конкретные ситуации приведены, например, в работе Т.Г. Воловой с соавторами (2008г.). При этом обычно используют метод наименьших квадратов. Затем проводят проверку адекватности полученной модели. Только после этого ее можно применять для прогноза и подбора контролируемых воздействий на процесс. Пример будет приведен в конце раздела.

Влияние концентрации основного продукта метаболизма

Рост микроорганизмов часто зависит не только от концентрации субстрата s в биореакторе, но также и от p – концентрации продукта метаболизма. Причем обычно накопление продуктов снижает (ингибирует) скорость роста. Так, в процессе образования этанола дрожжами при накоплении спирта в культуре начинает нарастать процесс ингибирования клеток.

Ингибирование учитывается моделями различных типов, подбираемыми по данным экспериментов. Наиболее простой является **модель Хиншельвуда**, которая описывается уравнением:

$$\mu = \mu_m - Kp.$$

Графическое выражение этого уравнения дано на рис. 16.

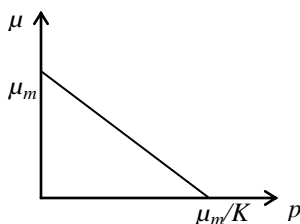


Рис. 16. Ингибирование удельной скорости роста микроорганизмов продуктом метаболизма по модели Хиншельвуда.

Более сложная **модель Иерусалимского** (рис. 17) описывается уравнением:

$$\mu = \frac{\mu_m K_p}{p + K_p},$$

где K_p - константа ингибирования продуктом, μ_m - максимальная удельная скорость роста.

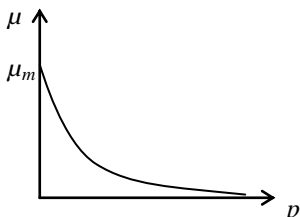


Рис. 17. Зависимость $\mu(p)$ по модели Иерусалимского

Модель частично ингибирующего продукта. Бывают ситуации, когда продукт частично ингибирует скорость роста микроорганизмов – не до нуля, а до определенного нижнего порога μ_0 . Эта ситуация графически выражена на рис. 18 (сравните с рис. 17).

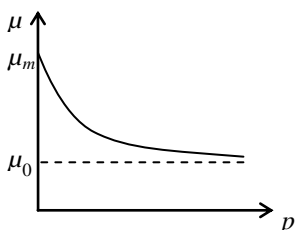


Рис. 18. Влияние частично ингибирующего продукта на удельную скорость роста микроорганизмов.

Модель стимулирующего продукта. Изредка встречаются процессы, в которых выделяемый клетками продукт метаболизма не ингибирует, а стимулирует рост культуры. Продукт является, по существу, стимулирующим субстратом или ферментом. Эта ситуация отражена на рис. 19.

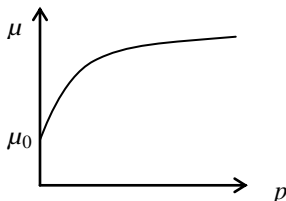


Рис. 19. Влияние стимулирующего продукта на удельную скорость роста микроорганизмов.

Влияние отмирания микроорганизмов

Все рассмотренные ранее модели касались собственно размножения микроорганизмов. Однако существует представление о том, что в процессе развития любой популяции одновременно с ростом происходит диссимиляция, т.е. отмирание микроорганизмов. Данное представление хорошо согласуется с экспериментальными данными биотехнологии. Поэтому модели роста биомассы должны учитывать влияние на μ не только b – удельной скорости рождаемости (размножения), но и d – удельной скорости смертности (диссимиляции) микроорганизмов:

$$\frac{dx}{dt} = (b - d)x. \quad (7)$$

Кроме простейших предположений о слабом влиянии диссимиляции ($d=0$) или о постоянном значении $d=K_d$, в уравнениях кинетики биомассы используют ряд более сложных моделей:

- **модель Ферхюльста**, где удельная скорость отмирания биомассы принята пропорциональной плотности (биомассе) популяции: $d=K_d x$;

- **модель Рамкришны**, где отражена кинетика взаимодействия продукта и биомассы микроорганизмов, а именно, удельная скорость диссимиляции биомассы пропорциональна p – концентрации продукта метаболизма: $d=K_d p$;

- **модель Колпикова**, которая связывает удельную скорость диссимиляции с влиянием s – концентрации субстрата:

$$d = \frac{\mu_d}{1 + s / K_d},$$

где μ_d – максимальная удельная скорость диссимилиации при нулевой концентрации субстрата, K_d – константа ингибирования процесса диссимилиации.

Эта модель дает переменную скорость отмирания в ходе процесса. Пока субстрата много, идет рост, а отмирания почти нет. С уменьшением концентрации субстрата скорость отмирания биомассы плавно повышается. Такая картина считается правдоподобной.

Для иллюстрации совместного влияния значений b и d непосредственно на $x(t)$ – плотность (биомассу) используем модель Ферхольста: $d=K_d x$. Уравнение (7) принимает вид, являющийся частным случаем уравнения Бернулли:

$$\frac{dx}{dt} = (b - K_d x)x.$$

Решение этого уравнения: $x(t)=(Ce^{-bt}+K_d/b)^{-1}$, где $C=x(0)^{-1}-K_d/b$.
 Ход кривой $x(t)$ существенно зависит от соотношения $x(0)$ – начальной плотности (биомассы) и K_d/b . Если $x(0)>(K_d/b)^{-1}$, то плотность постепенно снижается до уровня, равного $(K_d/b)^{-1}$ (кривая а на рис. 20), если $x(0)<(K_d/b)^{-1}$ – наоборот повышается до того же уровня (кривая б).

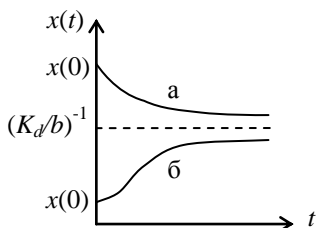


Рис. 20. Характер изменения плотности (биомассы) популяции. Пояснения в тексте

Помимо указанных факторов на кинетику роста микроорганизмов влияет температура, рН среды и т.п. Как и в любой кинетике, эти параметры оказывают влияние на константы в кинетических моделях роста микроорганизмов. В принципе любая из констант не является постоянной т.к. в той или иной мере может быть подвержена их влиянию. Однако в кинетических уравнениях для роста популяции чаще всего полагают, что, например, температура существенно влияет на μ_m – максимальную удельную скорость роста, а на остальные константы – в меньшей степени. Для учета подобных

влияний в приведенных выше моделях подбирают значения констант, соответствующих конкретным условиям процесса, или вносят функции-зависимости констант от условий среды в явном виде. Затем, как уже отмечалось, полученные экспериментальные данные обрабатывают методом наименьших квадратов и проверяют адекватность полученной модели.

3.3. Моделирование кинетики образования продукта метаболизма

Одновременно с ростом микроорганизмов или при определенном сдвиге во времени, но в том же биотехнологическом процессе происходит биосинтез продуктов метаболизма. Закономерности этого процесса также требуют своего математического описания. По аналогии с μ – удельной скоростью роста микроорганизмов вводится параметр q – **удельная скорость биосинтеза основного продукта**, которая входит в уравнение, аналогичное по структуре (3):

$$\frac{dp}{dt} = qx. \quad (8)$$

В моделях биотехнологических процессов параметр q для продуктов, связанных с ростом микроорганизмов, может быть выражен простой зависимостью $q=K\mu$. Более сложное выражение было предложено Людекингом и Пайри:

$$q = q_0 + K\mu.$$

В этом случае часть продукта образуется непрерывно, независимо от роста микроорганизмов, а часть – пропорционально удельной скорости роста популяции.

Используют и более сложные нелинейные зависимости q от μ :

$$q = \frac{a\mu}{b + \mu} \quad \text{или} \quad q = \frac{a\mu}{b - \mu}.$$

Первая из них дает выпуклую кривую, выходящую из начала координат (рис. 21 а), вторая – вогнутую (рис. 21 б).

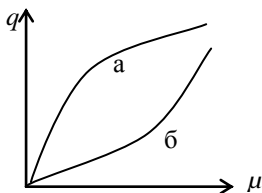


Рис. 21. Нелинейные зависимости q от μ . Пояснения в тексте

Исследования показали, что применимы практически такие же формы зависимостей q от p , s , как в моделях Моно, Мозера, Андрюса, Контуа и т.п. для μ – удельной скорости роста микроорганизмов. То же относится к влиянию на q температуры и pH среды. Идентификация и настройка модельных зависимостей для конкретных ситуаций, как уже отмечалось, производится эмпирически.

Для биосинтеза продуктов метаболизма часто бывает недостаточно учета названных факторов среды. Ведь в биосинтезе участвуют внутриклеточные ферменты микроорганизмов, промежуточные продукты, содержание которых в клетке зависит от предыстории развития культуры. Однако эти зависимости весьма сложны. Поэтому при моделировании используются различные упрощенные подходы для оценки влияния на q **возрастного состояния популяции микроорганизмов**.

В частности, японским ученым Аибой был предложен простой подход: использовать для оценки возраста культуры так называемый **средний возраст** как параметр, определяющий биосинтетическую активность культуры. Популяция условно делится на n возрастных групп. Θ_i обозначает возраст i -ой группы ($i=1..n$). Оценивается Δ_i - доля каждой группы в популяции, причем можно учесть изменения долей Δ_i от времени протекания биотехнологического процесса. Тогда средний, точнее средневзвешенный возраст микроорганизмов (клеток), участвующих в биосинтезе равен:

$$\bar{\Theta} = \sum_{i=1}^n \Theta_i \times \Delta_i .$$

Например, достаточно часто влияние параметра $\bar{\Theta}$ на кинетику биосинтеза продуктов метаболизма имеет возрастающий характер по q , но с насыщением (рис. 22). Такое влияние удобно выразить в форме, похожей на уравнение Моно:

$$q = \frac{q_m \bar{\Theta}}{K_p + \bar{\Theta}}, \quad \text{или} \quad q = q_0 + \frac{q_m \bar{\Theta}}{K_p + \bar{\Theta}} .$$

Первая из двух модельных кривых (рис. 22 кривая а) выходит из начала координат, вторая (рис. 22 кривая б) – из точки q_0 .

Если, наоборот, q убывает со средним возрастом, то лучше подходит выражение, подобное модели Иерусалимского (рис. 22 в, г):

$$q = \frac{q_m}{1 + \bar{\Theta} / K_p} \quad (\text{кривая в}) \quad \text{или} \quad q = q_0 + \frac{q_m}{1 + \bar{\Theta} / K_p} \quad (\text{кривая г}).$$

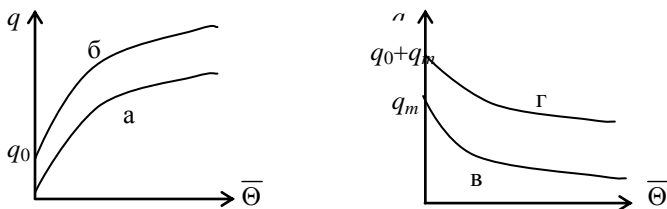


Рис. 22. Варианты модельных зависимостей q – удельной скорости биосинтеза продукта от среднего возраста микроорганизмов популяции. Пояснения в тексте.

Используются и другие модельные зависимости q от $\bar{\Theta}$. В частности, удачной во многих случаях оказывается кусочно-линейная аппроксимация с областью $(\bar{\Theta}_1 - \bar{\Theta}_2)$ максимума q (рис. 23).

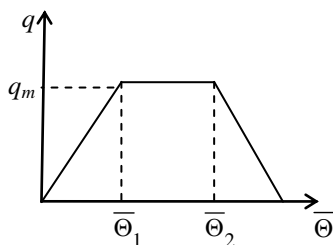


Рис. 23. Вариант кусочно-линейной зависимости q от $\bar{\Theta}$

В ряде случаев складывается ситуация, когда синтезированные продукты метаболизма остаются не совсем устойчивыми. Часто они настолько лабильны, что разрушаются уже в самом процессе синтеза. Поэтому, описывая материальный баланс по продукту метаболизма, необходимо учитывать кинетику его **инактивации** (деградации). Для этого в правую часть уравнения (8) добавляют член (Q) , отражающий общее снижение скорости изменения p – концентрации продукта из-за его инактивации:

$$\frac{dp}{dt} = qx - Q.$$

Возможны различные варианты моделирования кинетики деградации продукта:

- а) деградация отсутствует: $Q=0$;
- б) деградация идет с постоянной скоростью: $Q=K$;

в) скорость деградации продукта прямо пропорциональна его концентрации: $Q=Kp$;

г) деградация идет по уравнению т.н. химической кинетики с показателем степени n (большим или меньшим 1): $Q=Kp^n$;

д) деградация зависит не только от концентрации продукта, но и от плотности популяции: $Q=Kpx$;

е) скорость реакции деградации зависит от плотности популяции и возрастает с концентрацией продукта до предела, зависящего от биомассы:

$$Q = \frac{Q_m P x}{K + p}.$$

3.4. Моделирование кинетики утилизации субстрата.

Для полноты «баланса» частей биотехнологического процесса следует дополнить общую модель уравнением потребления субстрата, точнее скорости снижения содержания субстрата:

$$-\frac{ds}{dt} = Q_1 + Q_2 + Q_3. \quad (9)$$

В правой части суммируются затраты на собственно рост популяции микроорганизмов (Q_1), на образование продукта метаболизма (Q_2) и на поддержание жизнедеятельности микроорганизмов (Q_3). Q_1 в простейшем варианте связано с правой частью уравнения (3): $Q_1=K_1\mu x$, где K_1 – константа. Q_2 можно выразить через правую часть уравнения (8): $Q_2=K_2qx$. Q_3 упрощенно можно считать величиной, пропорциональной плотности микроорганизмов: $Q_3=K_3x$.

3.5. Моделирование накопления L-лейцина.

L-лейцин – незаменимая аминокислота, необходимая для промышленного получения лизина, т.к. производство последнего базируется на лейцинозависимых штаммах. Лизин широко применяют в сельском хозяйстве как кормовая добавка. Сам L-лейцин применяется, например, в спортивном питании, т.к. является предшественником незаменимых жирных кислот, входящих в состав клеточных мембран. Основной способ производства L-лейцина – микробиологический синтез с использованием штамма *Corynebacterium glutamicum*. Рассмотрим пример моделирования процесса получения L-лейцина из работы Д.С. Осипова (2002г.).

Пример. Биосинтез проводился в лабораторном биореакторе. Для построения модели в отбираемых пробах определялась оптическая плотность раствора – x , содержание L-лейцина – p и содержание субстрат-редуцирующих веществ (РВ) по Бертрану – s . По ходу процесса биосинтеза (время t) были получены пять экспериментальных измерений этих показателей. Результаты измерений основных показателей в процессе накопления продукта метаболизма – L-лейцина приводятся в таблице.

Продолжительность процесса (час)	$t=0$	$t=6$	$t=12$	$t=19$	$t=26$
Оптическая плотность раствора (x), ед	0,020	0,087	0,375	0,705	0,895
Концентрация L-лейцина (p), г/л	0	1,8	6,3	16,2	25,2
Δs – потребление РВ, г/л	0	6,2	48,0	112,5	186,9

Для описания динамики показателя x – оптической плотности, отражающей биомассу штамма, использовалось уравнение (3) с правой частью, соответствующей модели Моно:

$$\mu = \frac{\mu_m s}{K_s + s}, \quad (10)$$

Модельное предположение о механизме биосинтеза продукта (L-лейцина) состоит в том, что с одной стороны, он осуществляется всеми покоящимися клетками (биомассой x), а с другой ассоциирован с ростом популяции (dx/dt). Поэтому для моделирования динамики показателя p было применено следующее дополненное соотношение (8):

$$\frac{dp}{dt} = ax + b \frac{dx}{dt}. \quad (11)$$

Далее, предполагается, что в соответствии с уравнением (9) субстрат расходуется:

- на синтез биомассы штамма ($Q_1=K_1\mu x$),
- на синтез L-лейцина: $Q_2=K_2(ax + b \frac{dx}{dt})$,
- на поддержание жизнедеятельности популяции ($Q_3=K_3x$).

Подставляя выражения Q_1 , Q_2 , Q_3 в уравнение (9), после упрощения получаем уравнение для расхода субстрата:

$$-\frac{ds}{dt} = c \frac{dx}{dt} + fx.$$

Все эмпирические константы в уравнениях, описывающих кинетику этого процесса, оценивали с помощью метода наименьших квадратов (МНК) на основе экспериментальных результатов измерений из предыдущей таблицы. Например, в последнем уравнении наилучшими оказались следующие МНК-оценки констант: $c=81,83$, $f=10,51$. Среднеквадратическая погрешность показателя s составила $S_s=\pm 1,37$.

В следующей таблицы приведены оценки основных показателей процесса, полученные в результате решения уравнений модели биосинтеза L-лейцина.

Продолжительность процесса (час)	$t=0$	$t=6$	$t=12$	$t=19$	$t=26$
Оптическая плотность раствора, ед	0,000	0,104	0,353	0,703	0,897
Концентрация L-лейцина, г/л	0	1,2	6,9	15,8	25,3
потребление РВ, г/л	0	8,1	47,4	112,6	186,9

Сравнивая экспериментальные данные измерений с этими оценками по модели можно сделать вывод, что уравнения в данном случае достаточно хорошо описывают биосинтез продукта метаболизма. Дополнительные статистические проверки показали адекватность такой модели.

Заключение

Таким образом, простейшую дескриптивную модель кинетики биотехнологического процесса можно представить системой из трех дифференциальных уравнений: (3), (8) и (9) с различными вариантами правых частей, рассмотренными в предыдущих разделах. Создание адекватной модели конкретного процесса требует выбора формы зависимости для каждой правой части трех уравнений (идентификация уравнений). Совокупность зависимостей дает модель, вернее структуру модели, потому что дополнительно необходимо подобрать значения констант-коэффициентов, входящих в уравнения (настройка модели).

Кроме того, рассмотренные кинетические зависимости следует дополнить математическими моделями массообмена, макро- и микросмещения, теплообмена, гидродинамики и других процессов, протекающих в ходе общего биотехнологического процесса. Но даже после этого в моделях не учитывается влияние управляемых воздействий на процесс, таких как пополнение субстрата, промежуточное частичное извлечение или пополнение биомассы, продукта и пр. Для этих целей разработаны более сложные варианты моделей, которые можно найти, например, в работе И.И. Протопопова и Ф.Ф. Пашенко (2004г.).

Контрольные вопросы. 1. Каковы основные количественные показатели, используемые при моделировании кинетики биотехнологических процессов? 2. Привести графики простейших зависимостей удельной скорости роста биомассы (плотности) популяции от концентрации основного компонента субстрата, а также от концентрации продукта метаболизма. Сформулировать модельные предположения для каждого графика. 3. Задать конкретные значения констант в модели Колпикова, которая предполагает ингибирующее влияние концентрации субстрата на популяцию микроорганизмов. Построить соответствующий график зависимости удельной скорости диссимилиации микроорганизмов от s . 4. По аналогии с моделями Моно, Мозеса и Андрюса составить формулы зависимости удельной скорости биосинтеза основного продукта биотехнологического процесса от s . Пояснить смысл каждой зависимости. 5. Пояснить способ оценки среднего возраста культуры в биореакторе. Какие модельные зависимости предложены для описания влияния среднего возраста на удельную скорость биосинтеза продукта? 6. Каковы модельные предположения о характере деградации продукта метаболизма в процессе его биосинтеза? 8. Самостоятельно изобразить графики зависимостей a - e , приведенных в конце раздела 3.3. 7. Какова структура уравнения расхода субстрата в биотехнологическом процессе?

4. Вероятностные модели

Все рассмотренные выше модели были **детерминистическими**, то есть в них не учитывались случайности и их влияние на изучаемые процессы, например, на численность популяции.

Существуют несколько причин, по которым детерминистические модели не всегда служат достаточно точным отражением реальности в биологии и в других областях знаний. Во-первых, они предполагают большую численность популяции. Настолько большую, что можно опираться на закон больших чисел, который включает в себя несколько фундаментальных теорем.

Одно из следствий этого закона используется в генетике: при достаточно большом объеме выборки из большой популяции F_2 , полученной из F_1 ($AA \times aa \rightarrow F_1$), соотношение особей с доминантным и рецессивным проявлениями признака близко к 3:1. Это удачная детерминистическая модель – закон Менделя, который, кроме прочих предположений, требует большой объем выборки в F_2 . В реальных экспериментах, всегда ограниченных по объему, уже по этой причине наблюдается отклонение от модели 3:1.

Теория вероятностей и основанная на ней математическая статистика представляют не только упрощенные детерминистические, но и так называемые **стохастические** варианты моделей и методов. Они, в частности, позволяют ответить на вопрос: можно ли считать отклонение от детерминистической модели 3:1, обнаруженное в ограниченной выборке, чисто случайным, естественным или, всё-таки, отклонение вызвано другими генетическими причинами (отбор, миграция и т.п.). Ответ всегда дается в вероятностной форме. Например, с помощью известного критерия χ^2 можно получить вывод: с вероятностью P менее 0,05 (5%) обнаруженное в эксперименте отклонение от 3:1 вызвано случайной ошибкой выборочности. Значит отклонение в действительности вызвано какими-то особыми, неслучайными факторами.

Вторая причина, по которой необходимо применять стохастические модели, особенно в сельском хозяйстве, это влияние на моделируемые процессы (системы), например, растения, неконтролируемых случайных колебаний условий выращивания. Часто влияние этих колебаний даже перекрывает влияние на урожай самих генотипов сравниваемых сортов. Для селекционеров и сортоиспытателей крайне важно грамотно применять стохастические модели и методы. Эти модели используют также в генетике и теории эволюции.

Пример. Оценим вероятность случайной мутации гена, который определяет группу крови. Пусть этот ген имеет три

аллеля (B, C, E). Вероятности спонтанных мутаций отдельных нуклеотидов: $A \leftrightarrow T$; $C \leftrightarrow G$ и т. д. одинаковы и приблизительно равны 10^{-2} . Аллель B отличается от C всего по десяти нуклеотидам, а «промежуточные» мутации нежизнеспособны. Если случайные замены нуклеотидов возникают независимо, то вероятность мутации от B к C т.е. $P(B \rightarrow C)$ приблизительно равна 10^{-20} . Для реализации такой ничтожно малой вероятности нужны огромные популяции и очень длинная череда поколений. Значит, если новый вид содержит все три аллеля (B, C, E), то они были получены от вида – предка, а не образовались в результате новых мутаций.

4.1. Сумма и произведение событий

Напомним, что в теории вероятности значок «+» между двумя событиями (A, B) означает «или», а «·» означает «и». Если события A и B несовместные, то $P(A+B)=P(A)+P(B)$. Если A и B независимы, то $P(A \cdot B)=P(A) \cdot P(B)$

Пользуясь только этими формулами можно доказать справедливость закона Менделя (1:2:1) в детерминистической форме.

Рассмотрим гибрид F_1 ржи (Rr). Этот гибрид образует мужские и женские гаметы:

Гаметы	♀ R	♀ r	♂ R	♂ r
Событие	A	B	C	D

Вероятности (доли, частоты) образования этих гамет равны $P(A)=P(B)=1/2$, $P(C)=P(D)=1/2$.

В процессе перекрёстного опыления с образованием семян F_2 происходит случайное объединение гамет. Возможны 4 варианта объединения:

Вариант объединения гамет (сложного события)	Вероятность объединения	Генотип F_2
A и C	$P(A \cdot C)=1/2 \cdot 1/2=1/4$	RR
A и D	$P(A \cdot D)=1/2 \cdot 1/2=1/4$	Rr
B и C	$P(B \cdot C)=1/2 \cdot 1/2=1/4$	rR
B и D	$P(B \cdot D)=1/2 \cdot 1/2=1/4$	rr

Эти вероятности случайного объединения гамет проще оценить с использованием решетки Пеннета.

♀ \ ♂	$R(1/2)$	$r(1/2)$
$R(1/2)$	$RR(1/4)$	$Rr(1/4)$
$r(1/2)$	$rR(1/4)$	$rr(1/4)$

Поскольку генотип Rr то же, что и rR , то сумма вероятностей появления этих генотипов в F_2 : $P(A \cdot D) + P(B \cdot C) = 1/4 + 1/4 = 1/2$.

Теперь можно определить соотношение генотипов в F_2 .

Генотип F_2	RR	Rr	rr
Вероятность образования	$1/4$	$1/2$	$1/4$
Соотношение	1	2	1

В процессе доказательства было использовано тождество вероятностей и частот (долей) образования гамет (R , r), а также генотипов (RR , Rr , rr). Это справедливо, так как по теореме Бернулли (одна из теорем закона больших чисел) при увеличении числа испытаний (объёма выборки растений популяции, случайно взятых гамет и зигот – семян F_2) доли семян с генотипами RR , Rr , rr в F_2 приближаются к вероятностям их образования, а значит, их соотношение к $1:2:1$. Таким образом, при увеличении выборки (объёма) популяции F_2 получаем закон Менделя для соотношения долей. То есть при увеличении выборки стохастическая модель превращается в детерминистическую модель для долей.

Рассмотрим некоторые простые формулы теории вероятности, которые при решении задач из генетики и селекции могут привести к весьма полезным для практики приложениям теории вероятности.

Если случайные события A_1, \dots, A_n происходят независимо друг от друга, то

$$P(A_1 \cdot A_2 \cdot A_3 \cdot \dots \cdot A_n) = \prod_{i=1}^n P(A_i),$$

где \prod – знак произведения.

Поэтому, если проводится n одинаковых независимых испытаний и в каждом вероятность события B равна p , то вероятность того, что событие B ни разу не произошло равна $(1 - p)(1 - p) \dots (1 - p) =$

q^n , где $q = 1 - p$. Вероятность противоположного события: «хотя бы один раз событие B произошло» равна: $1 - q^n$.

Рассмотрим пример. Оценить вероятность того, что среди 8 особей потомства F_2 от скрещивания белой (cc) и серой (CC) мыши будет хотя бы одна белая (C – доминантный аллель).

Сначала оценим вероятность того, что в F_2 не будет ни одной белой мыши. Так как доля серых мышей в F_2 составляет $\frac{3}{4}$ ($CC + Cc$), то искомая вероятность будет $(\frac{3}{4})^8 \approx 0,1$. Теперь можно определить вероятность того, что будет хотя бы одна белая: $1 - 0,1 = 0,9$. Эта стохастическая модель имеет и детерминистическую формулировку: среди большого числа семей F_2 с 8 потомками в 90% семей будет хотя бы одна белая мышь (в 10% таких семей – ни одной белой).

Задача 1. Из литературных данных известно, что вероятность успешной трансформации единичного растительного образца определенного вида чужеродным геном устойчивости к заболеванию с использованием агробактерии *A. tumefaciens* равна приблизительно 0,02. Достаточно ли провести эксперимент со ста однотипными эксплантами, чтобы с вероятностью не менее 0,95 получить хотя бы одно трансгенное растение?

Схема решения:

$(1 - p)^n = (1 - 0,02)^{100} \approx 0,13$ – это вероятность того, что из 100 эксплантов не удалось получить ни одного трансгенного растения; $1 - (1 - p)^n \approx 0,87$ – хотя бы одно трансгенное растение получено. Оценка этой вероятности меньше 0,95; значит 100 эксплантов не достаточно. Из уравнения $1 - (1 - 0,02)^n = 0,95$ следует, что для достижения вероятности 0,95 требуется, как минимум, $n \approx 150$ эксплантов.

Задача 2. Частота спонтанной мутации - альбинизма у растений 10^{-5} . Сколько семян надо высеять, чтобы с вероятностью 0,95 среди них было хотя бы одно альбиносное растение?

Схема решения:

$(1 - p)^n$ – вероятность того, что нет ни одного мутанта среди n растений; $1 - (1 - p)^n = 0,95$ – хотя бы одно мутантное растение есть. Из последнего равенства при $p = 10^{-5}$ можно найти n . В данном случае оно приблизительно равно 300 000.

Задача 3. Вероятность рождения мальчика и девочки равны $p = q = 1/2$. Сколько нужно планировать детей в семье,

чтобы вероятность иметь хотя бы одного мальчика была более 0,9?

Схема решения:

$P = (1/2)^n$ – из n детей все девочки; $1 - (1/2)^n > 0,9$ – хотя бы один мальчик. Отсюда $(1/2)^n < 0,1$.

Это неравенство справедливо уже при $n = 4$.

В качестве более сложного примера применения подобных вероятностных моделей можно предложить подход к оценке объёма выборки семян в семеноводстве.

Пусть сорт самоопыляющейся культуры – это смесь из N чистых линий (биотипов), содержащихся в сорте с известными долями p_i :

$$\sum_{i=1}^N p_i = 1$$

Как определить минимальный объём случайной выборки семян (n), чтобы в ней с вероятностью $P \geq 0,95$ было хотя бы по одному зерну каждого из N биотипов.

Итак, n – объём выборки, т.е. число случайно выбранных семян (испытаний); N – число биотипов;

$(1 - p_i)^n$ – вероятность того, что из n испытаний случайно не выберем ни одного семени i -го биотипа;

$1 - (1 - p_i)^n$ – вероятность того, что из n испытаний возьмем хотя бы одно семя i -го биотипа;

$$P(A) = \prod_{i=1}^N [1 - (1 - p_i)^n] - \text{вероятность того, что в ходе } n \text{ испытаний}$$

случайно возьмём хотя бы одно семя от каждого биотипа. Далее подбираем n – объём выборки семян. Не вдаваясь в подробности отметить, что эта формула верна, если n значительно (в 5 и более раз) превышает N .

4.2. Формула полной вероятности

Пусть в опыте событие A может наступить только вместе с одним из n несовместных событий B_1, B_2, \dots, B_n , составляющих полную группу событий [то есть $P(B_1 + B_2 + \dots + B_n) = 1$]. Тогда так называемая безусловная или полная вероятность реализации события A равна:

$$P(A) = \sum_{i=1}^n P(B_i) \times P_{B_i}(A),$$

где $P_{B_i}(A)$ – условная вероятность наступления события A вместе с B_i .

Задача 1. Из потомства F_2 от скрещивания белой (cc) и серой (CC) мышей взяли одну серую и скрестили с белой (C – доминантный аллель). Какова вероятность того, что из трёх мышей – потомков такого скрещивания – все три серые?

Схема решения:

Результаты исходных скрещиваний до F_2 включительно удобно представить в следующем виде:

P:	cc	x	CC	
		↓		
F ₁		Cc	x	Cc
			↓	
F ₂		CC	Cc	cc
		1/4	1/2	1/4

Серая мышь, взятая из F_2 для скрещивания с белой не могла иметь генотип cc , а только CC (событие B_1) или Cc (событие B_2). Поэтому B_1 и B_2 – события, составляющие полную группу. Значит, вероятность суммы этих событий равна 1. “Внутри” этой суммы вероятность Cc в два раза больше, чем CC . Следовательно, вероятность того, что взятая из F_2 серая мышь имела генотип CC равна $P(B_1) = 1/3$, а генотип Cc $P(B_2) = 2/3$.

Если для скрещивания с белой мышью случайно взяли серую мышь с генотипом CC , то после скрещивания $CC \times cc$ в потомстве будут все серые мыши, значит для события B_1 условная вероятность $P_{B_1}(A)$ появления всех трёх серых мышей в потомстве равна 1^3 . Если же для скрещивания случайно взяли серую мышь с генотипом Cc (событие B_2), то после скрещивания $Cc \times cc$ в потомстве будет $1/2$ серых (Cc) и $1/2$ белых (cc) мышей, и условная вероятность $P_{B_2}(A)$ того, что все три потомка будут серыми, равна $(1/2)^3$.

Теперь можно оценить искомую полную вероятность $P(A) = 1/3 \cdot (1)^3 + 2/3 \cdot (1/2)^3 = 5/12 = 0,42$

Задача 2. Определим соотношение долей генотипов в F_3 после самоопыления популяции F_2 пшеницы, полученной из F_1 ($DD \times dd$).

Схема решения:

По формуле полной вероятности событие A “зерно, случайно взятое из F_3 , несет генотип DD ” имеет вероятность: $P(DD) = 1/4 \cdot 1 + 1/2 \cdot 1/4 + 1/4 \cdot 0 = 0,375$.

Действительно, событие A может произойти совместно с одним из 3-х событий:

1) B_1 – это зерно вызрело на растении F_2 , имеющем генотип DD ;

2) B_2 – это зерно вызрело на растении F_2 , имеющем генотип Dd ;

3) B_3 – это зерно вызрело на растении F_2 , имеющем генотип dd .

Соотношение долей трех генотипов в F_2 : $1/4(DD)$, $1/2(Dd)$, $1/4(dd)$.

Поэтому вероятность того, что случайно взятое зерно вызрело на растении DD из F_2 : $P(B_1) = 1/4$;

$P_{B_1}(A) = 1$ – вероятность того, что это случайно взятое зерно, образовавшееся в результате самоопыления на растении DD , имеет генотип DD ;

$P(B_2) = 1/2$ – вероятность того, что зерно созрело на растении F_2 с генотипом Dd ;

$P_{B_2}(A) = 1/4$ – вероятность того, что это зерно с растения Dd имеет генотип DD ;

$P(B_3) = 1/4$ – вероятность того, что зерно взято с растения F_2 , несущего генотип dd ;

$P_{B_3}(A) = 0$ – вероятность того, что случайно взятое зерно с такого растения имеет генотип DD .

B_1, B_2, B_3 – полная группа событий.

Аналогично, вероятности случайно взять из F_3 зерно Dd и dd :

$$P(Dd) = 1/4 \cdot 0 + 1/2 \cdot 1/2 + 1/4 \cdot 0 = 0,25.$$

$$P(dd) = 0,375.$$

При большом объеме случайной выборки семян из F_3 три оцененные вероятности $P(DD)$, $P(Dd)$, $P(dd)$ близки к долям трёх генотипов. Таким образом, в F_3 будет следующее соотношение долей генотипов:

DD	Dd	dd
0,375	0,25	0,375

4.3. Теория мишени

В некоторых областях биологии удобно пользоваться так называемой теорией мишени. Согласно этой модели каждая клетка,

организм или ген представляют собой “мишень”, а происходящие с ними изменения – результат случайных ударов (попаданий) при бомбардировке этих мишеней. Такую модель применяют для оценки эффективности воздействия индуцированного ионизирующего излучения на живые организмы. Эта теория естественным образом была расширена и для объяснения “спонтанных” вредных изменений, в частности изменений, связанных с процессами старения, а также действия различных факторов помимо ионизирующего излучения. Те же самые математические идеи находят применение и во многих других областях биологии, в частности в экологии.

Предположим, что клетка содержит несколько мишеней (мутирующих генов) и попадание нейтрона в любую из них с известной вероятностью вызывает мутацию с заметным эффектом у потомства. Какая доля из клеток популяции после облучения определенной дозой (бомбардировка нейтронами) будет иметь хотя бы один мутантный ген и какая ни одного?

Примеры:

1. Сперма дрозофилы бомбардируется нейтронами. В хромосомах спермы имеется множество генов (“мишеней”), каждый из которых важен для нормального развития. Существуют методы скрещивания, позволяющие определить по потомству, в какой доле сперматозоидов хотя бы один из этих генов мутировал. Как изменяется эта доля сперматозоидов в зависимости от дозы?

2. Половозрелое гаплоидное насекомое (например, трутень) облучается рентгеновскими лучами. Предположим, что в каждой клетке содержится N генов, каждый из которых существенен для её нормального функционирования. Как зависит от дозы доля клеток, перестающая нормально функционировать в результате облучения, то есть доля клеток с явными мутациями?

В обоих примерах мишенями являются гены.

Уточним математическую модель для оценки доли клеток хотя бы с одной мутацией. Итак, каждая клетка содержит N мишеней и подвергается действию дозы в k частиц. Пусть вероятность того, что определенная частица “попадет в определенную мишень” (вызовет мутацию гена), равна p . Понятно, что p очень мало. Вероятность того, что данная мишень не будет поражена данной частицей, равна $1 - p$.

Следовательно, вероятность того, что данная мишень не будет поражена ни одной из k частиц, равна $(1 - p)^k$. Если p мало, а k – велико, то удобной для расчетов и достаточно точной является следующая замена: $(1 - p)^k \approx e^{-kp}$.

В клетке имеется N мишеней, и вероятность поражения данной мишени не зависит от поражения остальных мишеней. Вероятность того, что этими k частицами не будет поражена ни одна из N мишеней клетки: $S = (e^{-kp})^N = e^{-Nkp}$. При большом числе клеток эта вероятность близка к доле клеток без мутаций. Тогда зависимость $1-S$ – доли клеток, несущих хотя бы по одной мутации, от k – дозы облучения – имеет вид $1 - e^{-Nkp}$ (сплошная линия на рис. 24).

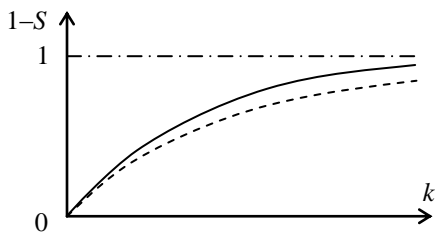


Рис. 24. Зависимость доли мутантных клеток ($1-S$) от дозы (k)

Следует напомнить, что любая модель имеет ограничения. В частности, полученная зависимость предполагает равную выживаемость клеток. Если же в действительности до оценки доли клеток без мутаций происходит частичная гибель мутантных клеток, то экспериментальная кривая, скорее всего, отклонится вниз по сравнению с прогнозом по модели (пунктир на рисунке). В подобных случаях следует усложнять модели – учитывать в них дополнительные гипотезы о биологических процессах. Адекватность новой модели косвенно подтвердит справедливость и достаточность всей совокупности гипотез – предположений.

Ряд Пуассона

Производится серия однотипных испытаний. Любое испытание успешно с вероятностью p и неуспешно с вероятностью $q = 1 - p$. Необходимо оценить вероятность сложного события A , состоящего в том, что из n испытаний k прошли успешно. Для этого, как известно, используют биномиальное распределение:

$$P_n(k) = C_n^k p^k q^{n-k}, \text{ где } C_n^k = \frac{n!}{k!(n-k)!};$$

Пример. Соотношение полов 1:1. Определить вероятность того, что в семье с тремя детьми будут 2 девочки.

Решение. $n = 3, k = 2$

$$C_3^2 = \frac{3!}{2!(3-2)!} = 3;$$

$$P(A) = 3 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^1 = 3/8..$$

Рассмотрим видоизмененное биномиальное распределение, широко используемое в биологии. Итак, если вероятность данного события при однократном испытании равна p , то вероятности того, что в ряду n последовательных испытаний оно (попадание) произойдет 0, 1, 2, 3, и т.д. раз, задаются соответствующими членами ряда:

$$\begin{array}{ccccccc} \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} & \dots & & \\ q^n & + npq^{n-1} & + \frac{n(n-1)}{2!} p^2 q^{n-2} & + \frac{n(n-1)(n-2)}{3!} p^3 q^{n-3} & \dots & = & 1 \end{array}$$

Вынесем q^n за скобки

$$q^n \left[1 + n \frac{p}{q} + \frac{n(n-1)}{2!} \left(\frac{p}{q}\right)^2 + \frac{n(n-1)(n-2)}{3!} \left(\frac{p}{q}\right)^3 \dots \right] = 1$$

Предположим теперь, что p очень мало ($q \approx 1$), а n достаточно велико, но так, что величина $m = np$ не является пренебрежимо малой. Тогда можно считать, что $n-1 \approx n-2, \dots \approx n$ и, используя замену $q^n = (1-p)^n \approx e^{-np}$, последнее выражение можно переписать в виде:

$$e^{-np} \left[1 + np + \frac{1}{2!} (np)^2 + \frac{1}{3!} (np)^3 \dots \right]$$

Заметим, что сумма, стоящая в скобках, равна e^{np} , а всё выражение равно единице: $e^{-np} [e^{np}] = 1$. Напомним, что соответствующие члены ряда представляют собой вероятности того, что в последовательности из n испытаний интересующее нас событие произойдет 0, 1, 2, 3, ... раз. Сумма всех вероятностей, как и следовало ожидать, равна единице.

Этот ряд, удобный для приблизительных экспресс-оценок, называют рядом Пуассона. Его используют во многих областях биологии. Он применим везде, где можно представить себе длинную последовательность независимых испытаний с малой вероятностью “успеха” в каждом испытании. Интуитивно ясно, что среднее число успехов m в последовательности n испытаний равно np . Поэтому, ряд Пуассона даёт вероятности данного числа успехов в последовательности n испытаний, причём в форме, удобной для вычислений.

Число успехов	0	1	2 ...	r
Вероятность	e^{-m}	me^{-m}	$\frac{1}{2!}m^2e^{-m}$	$\frac{1}{r!}m^r e^{-m}$

Докажем, что среднее число успехов, как утверждается выше, равно $m = np$.

Математическое ожидание числа успехов равно:

$$0 \cdot e^{-m} + 1 \cdot me^{-m} + 2 \frac{1}{2!}m^2e^{-m} + \dots = me^{-m}(1 + m + \dots) = m = n \cdot p$$

При исследовании действия излучения моделирование в рамках теории мишени представляется вполне естественным. Однако этот же математический аппарат можно применять в задачах, в которых аналогия с мишенями и снарядами выражена менее явно.

Рассмотрим пример. Предположим, что большая популяция из N бактерий смешана с популяцией из kN фаговых частиц. Какова будет доля незараженных бактерий, если допустить, что “нападение” фаговой частицы на любую бактерию равновероятно и все фаговые частицы проникают в бактерии? Может быть поставлена и обратная задача: сколько вирусных частиц k приходилось в среднем на одну бактерию, если доля бактерий, оставшихся незараженными равна F ?

Если рассматривать бактерии как мишени, а фаговые частицы как снаряды, то вероятность того, что на данную бактерию нападет данная фаговая частица, равна $1/N$. При большом числе N доли бактерий, зараженных 0, 1, 2, ... фаговыми частицами, задаются членами ряда Пуассона при $m = np = kN(1/N) = k$:

$$e^{-k} \left(1 + k + \frac{k^2}{2!} + \frac{k^3}{3!} + \dots \right).$$

В частности, вероятность того, что данная бактерия вообще избежит заражения, равна e^{-k} . При большом N эта вероятность равна доле незараженных бактерий.

Аналогичная задача возникает при подсчёте клеток или других микрочастиц под микроскопом с помощью специальной сетки. Предположим, например, что капля крови, содержащая N эритроцитов, размазана по предметному стеклу, разделенному на 400 одинаковых квадратов. Если эритроциты распределены по стеклу случайно, вероятность того, что данный эритроцит попадет в определённый квадрат, равна $1/400$, и, следовательно, ожидаемые числа квадратов с 0, 1, 2, ... эритроцитами задаются членами ряда Пуассона.

В данном случае $m = n \cdot p = N/400$. Ожидаемая доля пустых клеток равна $e^{-N/400}$; ожидаемое их число $400 \cdot e^{-N/400}$.

Если, например, подсчитали, что в 61 из 400 квадратов нет ни одного эритроцита, то: $400 \cdot e^{-N/400} \approx 61$. Значит общее число эритроцитов в капле крови на всем стекле $N \approx 752$.

Итак, если известно, что эритроциты распределены случайно, то таким способом можно быстро определить их примерное количество, подсчитав под микроскопом лишь число пустых квадратов сетки. Учитывая ограниченный объем выборки эритроцитов, следует ожидать некоторую ошибку выборочности. Для уточнения оценки N можно подсчитать число клеток с одним эритроцитом. Например, получили 120 таких клеток сетки. По формуле ряда Пуассона: $N \cdot e^{-N/400} = 120$. Откуда $N \approx 768$. Усреднение двух оценок дает: $N \approx (752+768)/2 = 760$.

Часто задаются не вероятностью p и число испытаний n , а сразу характерное значение параметра $m = np$, то есть среднее значение наступления события A во всей большой серии испытаний. Это значение m находят заранее при статистической обработке данных.

Пример. В травматологическое отделение в течение каждого часа дневного времени привозят примерно 3 пациентов. Какова вероятность того, что за один час их будет только 1 или ни одного?

В этом случае $m=3$ и $P(1) = me^{-m} = 3e^{-3} = 0,15$, а $P(0) = e^{-m} = 0,05$. Таким образом, вероятность того, что в течение часа потребуется по крайней мере 1 хирург, равна $1-P(0)=0,95$, а что таких хирургов нужно будет не менее 2, равна $1-P(0)-P(1)=0,8$. Как видно, это отделение травматологии нельзя закрывать на обед.

Приложения в экологии

В экологии ряд Пуассона используют, в частности, для того, чтобы выяснить, действительно ли организмы на обследуемом участке распределены случайно. Для этого участок разбивают на много одинаковых квадратов и подсчитывают число особей данного вида растений или животных в каждом квадрате. Если участок слишком велик, то выбирают наугад несколько квадратов и в каждом из них подсчитывают число особей. В любом случае числа квадратов, на которых обнаружено 0, 1, 2, ..., k особей, сравнивают (используя критерий χ^2) с ожидаемыми числами при пуассоновском распределении, то есть в предположении, что организмы распределены по участку совершенно случайно.

Критерием χ^2 можно пользоваться, если соблюдаются определённые условия: достаточно большой объём выборочной совокупности (например, общее число пересчитанных животных $n > 50$);

в каждой выделенной группе ожидаемое число особей должно быть не менее пяти (например, число квадратов с тремя животными должно быть не менее пяти); для вычисления χ^2 используют только численности, а не доли, проценты или величины, полученные при измерениях или взвешиваниях и т.д.

Если наблюдаемые и ожидаемые значения совпадают довольно хорошо ($\chi^2_p < \chi^2_{т}$), то можно сделать вывод, что распределение организмов по участку близко к случайному (не путать с равномерным). В противном случае ($\chi^2_p > \chi^2_{т}$), есть две возможности:

1. Особи избегают друг друга или же препятствуют пребыванию поблизости от себя других особей. В таких ситуациях если, например, среднее число особей на квадрат равно трём, то квадратов с тремя особями будет слишком много по сравнению с ожидаемым числом. Квадратов же, в которых нет ни одной особи или, наоборот, много особей, будет слишком мало, так как они стремятся разместиться более равномерно по участку.

2. Особи “сгущаются”, например, потому, что они привлекают друг друга, или потому, что в некоторых местах рассматриваемого участка условия более благоприятны для их существования, чем на других. В этом случае будет слишком много как пустых квадратов, так и квадратов с большим количеством особей.

Задача. На некотором достаточно однородном участке было расставлено 543 ловушки разных типов (конструкций) для мелких млекопитающих. По прошествии некоторого времени в ловушках было обнаружено:

Число животных в ловушке	0	1	2	3	4	5	6	7
Число ловушек	468	41	16	11	2	4	0	1

Ловушек с 8-ю и большим числом животных не было.

Спрашивается – одинаковы ли по эффективности ловушки разных конструкций?

Схема решения:

Общее число пойманных животных:

$$n = 0 \cdot 468 + 1 \cdot 41 + 2 \cdot 16 + 3 \cdot 11 + 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 0 + 7 \cdot 1 = 141,$$

$$m = n \cdot p = 141 \cdot 1 / 543 = 0,26.$$

Теперь, используя ряд Пуассона, можно определить ожидаемые доли ловушек с 0, 1, 2, и т.д. животными, в предположении, что все ловушки равно эффективны. После чего находим ожидаемое число ловушек с определенным числом животных.

Ожидаемая доля ловушек, в которых нет ни одного животного, равна $e^{-m} = e^{-0,26} = 0,77$. Ожидаемое число ловушек, в которых нет ни одного животного $0,77 \cdot 543 \approx 418,1$.

Аналогично определяют ожидаемые доли и число ловушек с 1, 2, 3-мя животными и т.д. Ожидаемую долю ловушек с числом животных от 4 до 7 можно определить по формуле $1 - P(0) - P(1) - P(2) - P(3) = 1 - 0,77 - 0,20 - 0,026 - 0,002 = 0,002$.

Число животных в ловушке	0	1	2	3	от 4 до 7	Σ
Ожидаемая доля ловушек	0,77	0,20	0,026	0,002	0,002	1
Ожидаемое число ловушек	418,1	108,6	14,1	1,1	1,1	543

Поскольку ожидаемое число ловушек с тремя и более животными меньше пяти, то три последних класса следует объединить. Тогда в этом объединенном классе общее ожидаемое число ловушек (с 2-я и большим числом животных – в таблице отмечено жирным шрифтом) будет 16,3, а наблюдаемое – 34. Всего, таким образом, сформировалось 3 класса.

$$\chi^2_p \approx \frac{(468 - 418,1)^2}{418,1} + \frac{(41 - 108,6)^2}{108,6} + \frac{(34 - 16,3)^2}{16,3} \approx 67,25$$

Для $P = 0,05$ при числе степеней свободы $df = 3 - 1 = 2$ табличное значение $\chi^2_{\tau} = 5,99$. $\chi^2_p > \chi^2_{\tau}$. Следовательно, не все конструкции ловушек равны по эффективности.

Итак, случайным (пуассоновским) является распределение объектов (клеток на предметном стекле; сорняков в поле и т.д.) аналогичное тому, которое получается в результате следующего модельного процесса:

1. Рассматриваемую площадь делят на большое число равновеликих квадратов;

2. Объекты помещают один за другим на случайно, независимо выбираемые квадраты. То есть, вероятность выбора данного квадрата совершенно одинакова и не зависит от того, содержит ли уже этот квадрат ноль, один или несколько объектов.

Редкие болезни, редкие признаки

Многие болезни достаточно редки или становятся таковыми после принятия профилактических и лечебных мер. Однако даже при самых благоприятных условиях в больших популяциях все же встречается некоторое число больных редкими заболеваниями. Распределение Пуассона даёт вероятности таких событий в нормальной ситуации. Если в наблюдаемой популяции, например, в конкретном городе, больные встречаются значительно чаще, чем это прогнозируется рядом Пуассона для аналогичных городов всей страны, то это говорит о нарушении условий в данном городе, о необходимости выяснения причин, принять меры и т.д.

Пример. При введении вакцины против полиомиелита иммунитет создаётся в 99,99% случаев. Какова вероятность того, что из 10000 вакцинированных детей заболеет 1?

Число «испытаний» $n = 10000$. Вероятность заболеть («попадание в мишень») $p = 0,0001$. Поэтому $m = np = 1$, и по формуле Пуассона имеем $P(1) = 1 \cdot e^{-1} = 0,368$. Вероятность, что ни один ребенок не заболеет: $P(0) = e^{-1} = 0,368$.

Аналогично, вероятность, что заболеют 2 ребёнка

$P(2) = \frac{1}{2!} e^{-1} \approx 0,184$, вероятности заболевания 3, 4 и 5 детей равны

соответственно $P(3) = \frac{1}{3!} m^3 e^{-m} \approx 0,061$; $P(4) = \frac{1}{4!} m^4 e^{-m} \approx 0,015$;

$P(5) = \frac{1}{5!} m^5 e^{-m} \approx 0,003$. Вероятность того, что хотя бы один

ребёнок из 10000 заболеет равна $1 - P(0) = 1 - e^{-1} \approx 0,632$.

Если принять, что 10000 новорождённых – это годовая норма, скажем, обычного областного роддома, то примерно в 73% [$P(0)+P(1)$] таких домов полиомиелитом заболеет не более одного ребенка в год; в 18% – два ребёнка в год; в 6% – три и в 1,5% – четыре ребёнка в год. Если же в каком-то роддоме заболело более 5 детей – то это чрезвычайное происшествие. Вероятность такого события равна 0,001.

Аналогичные расчёты можно провести по детской смертности, врождённым аномалиям и признакам.

Пример. В среднем по стране 1 из 600 детей рождается с болезнью Дауна. Следовательно, в каждом микрорайоне, где проживает 3000 детей, в среднем будет $m = np = 3000 \times 1/600 = 5$ детей, страдающих такой болезнью. При этом вероятность

рождения 10 детей из 3000 с синдромом Дауна в микрорайоне равна:

$$P(10) = \frac{m^{10}}{10!} e^{-m} = \frac{5^{10}}{10!} e^{-5} \approx 0,018 .$$

Таким образом, подобных микрорайонов должно быть приблизительно 2 из 100.

Контрольные вопросы. 1. Чем отличаются стохастические модели от детерминистических?. Пояснить на примерах. 2. Определить соотношение долей генотипов Aa и aa в популяции F_3 , полученной после самоопыления популяции F_2 пшеницы со структурой $0,25AA$; $0,5Aa$; $0,25aa$. 3. Привести примеры генетических, микробиологических, экологических и медицинских экспериментов, при анализе которых может быть применена теория мишени. 4. Для каких целей в экологии можно использовать ряд Пуассона? Пояснить на примерах.

5. Исследование операций на основе оптимизационных моделей

В наше время, которое по справедливости называют эпохой научно-технической революции, наука уделяет всё больше внимания вопросам организации и управления объектами, процессами. Это касается не только промышленности, но и биологии, медицины, сельского хозяйства, экологии и т. д. От науки требуются рекомендации по оптимальному (разумному) управлению процессами. Прошли времена, когда правильное, эффективное управление находилось организаторами «на ощупь», методом «проб и ошибок». Сегодня для выработки такого управления требуется научный подход – слишком велики потери, связанные с ошибками.

Потребности практики вызвали к жизни специальные научные методы, которые удобно объединять под названием **исследование операций**. Под этим термином будем понимать **применение математических моделей и методов для обоснования решений во всех областях целенаправленной человеческой деятельности**.

Поясним, что понимается под «решением». Пусть планируется какое-то мероприятие, направленное к достижению определённой цели. У лица, организующего мероприятие, всегда имеется какая-то свобода выбора: можно организовать его тем или другим способом, например, подобрать образцы техники, которые будут применены, так или иначе распределить средства и т.д. **Решение - это и есть какой-то выбор из ряда возможностей, имеющихся у организатора.**

Исследование операций начинается тогда, когда для обоснования решений применяется та или другая модель и математический аппарат. Исследование операций – это своеобразное математическое «примеривание» будущих решений, позволяющее экономить время, силы и материальные средства, избегать серьёзных ошибок.

Впервые название «исследование операций» появилось в годы второй мировой войны, когда в вооружённых силах некоторых стран (США, Англия) были сформированы специальные группы научных работников (физиков, математиков, инженеров), в задачу которых входила подготовка проектов решений для командующих боевыми действиями. Эти решения касались, главным образом, боевого применения оружия и распределения сил и средств по различным объектам. В дальнейшем, исследование операций расширило область своих применений на самые разные области практики: промышленность, сельское хозяйство, строительство, торговля, здравоохранение, охрана природы и т. д.

Чтобы познакомиться со спецификой этого направления прикладной науки рассмотрим ряд типичных для неё задач.

Продажа сезонных товаров. Для реализации определенной массы сезонных товаров создается сеть временных торговых точек. При этом финансовые и другие ресурсы всегда ограничены. Требуется выбрать разумным образом: число точек, их размещение, товарные запасы и количество персонала на каждой из них так, чтобы обеспечить максимальную экономическую эффективность распродажи.

Медицинское обследование. Известно, что в каком-то районе обнаружены случаи опасного заболевания. С целью выявления заболевших (или носителей инфекции) организуется медицинское обследование района. На это выделены ограниченные материальные средства, оборудование, медицинский персонал. Требуется разработать такой план обследования (число медпунктов, их размещение, последовательность осмотров специалистами, виды анализов и т.д.), который позволит выявить, по возможности, максимальный процент заболевших и носителей инфекции.

Селекционно-генетические исследования. В распоряжении имеется коллекция образцов растений, несущих картированные гены различных признаков: по одному, два, три и более генов (возможно сцепленных) в одном образце. Требуется разработать оптимальную схему выбора части образцов и их скрещиваний, чередующихся с отборами по фенотипу. Цель – за минимальное время вывести новый образец с заранее заданным новым сочетанием генов.

Деятельность сельскохозяйственного предприятия. Как лучше всего организовать деятельность крупного фермерского хозяйства – какие культуры и на каких площадях выращивать, в какой пропорции следует выделять средства для животноводства, птицеводства и т.д.

С чего начинать исследование? Прежде всего, нужно четко выделить факторы, которые существенно влияют на принимаемые решения. В последнем случае к ним относятся: количество земли, имеющейся в распоряжении хозяйства, ожидаемые урожайности культур, которые можно возделывать, ресурсы для создания животноводческой и птицеводческой продукции (помещения, корм и т. п.), а также ожидаемые потребности рынка в зерне, мясе, молоке, яйцах и многие, многие другие факторы. Некоторыми из них можно управлять в определенных пределах, остальные от нас не зависят.

Ясно, что, прежде всего, нужно среди управляемых выделить несколько **главных факторов**, возможно разбив общую деятельность на отдельные блоки. Это само по себе сложно сделать. Опыт и знания, накопленные ранее людьми, позволят выделить главные факторы,

влияющие на результат. В этом могут также помочь специальные математические методы.

Допустим, что так или иначе мы выделим существенные факторы. Что делать дальше? Теперь следует описать, каким же образом сказывается влияние этих факторов. Например, расширение помещений для скота позволяет увеличить численность стада. Чем выше урожайность какой-либо культуры (зависящая от сортов и элементов агротехники), тем больше дохода может быть получено от ее возделывания и т.п. Иными словами, дается качественная оценка факторов. Этого было достаточно для изучения задач с малым числом существенных факторов. К концу XX века положение резко изменилось. Современное высокоразвитое хозяйство требует более точных экономических рекомендаций. Уже мало сказать, например, что изменение фактора A на 1% даст прирост дохода на 1000 руб., если остальные факторы останутся неизменными. А если они все изменятся, что будет тогда? Может быть, эффект будет еще больше?

Чтобы ответить на эти вопросы, и создаётся **математическая модель управляемого объекта**, выражающая количественные соотношения между существенными факторами (параметрами) и, если снова вернуться к хозяйству, количественным выражением для оценки его деятельности – **целевая функция** (например, доход за год). Здесь и начинается, собственно, исследование операций.

Операцией называется всякое мероприятие (система действий), объединенное единым замыслом, и направленное к достижению какой-то цели (все мероприятия, рассмотренные выше, являются «операциями»). Исследование операций ведётся на модели.

Операция есть всегда управляемое мероприятие, то есть от нас зависит, каким способом выбрать некоторые параметры, характеризующие её организацию. «Организация» здесь понимается в широком смысле слова, включая набор технических, финансовых и других средств, применяемых в операции.

Всякий определённый выбор зависящих от нас параметров называется решением. Решения могут быть удачными и неудачными, разумными и неразумными. **Оптимальными называются решения, по тем или другим признакам предпочтительные перед другими. Цель исследования операций – предварительное количественное обоснование оптимальных решений.**

Иногда (относительно редко) в результате исследования удаётся указать одно – единственное строго оптимальное решение. Гораздо чаще выделяется целый ряд практически равноценных оптимальных решений – рекомендаций. Окончательный выбор решения всегда осуществляет человек.

5.1. Линейное программирование

Пусть из различных видов сырья, имеющегося в количествах, равных соответственно b_1, b_2, \dots, b_m (всего m видов сырья: например, сортов муки, специй и т.д.), может быть изготовлено n видов продуктов (например, n видов хлеба). Цена единицы j -го вида продукта на рынке равна c_j . Для получения единицы j -го продукта необходимо затратить i -й вид сырья в количестве a_{ij} единиц. Какие виды продуктов выгоднее всего изготавливать и сколько?

Прежде всего, нужно выяснить, в каком смысле понимаются слова «выгоднее всего». Так как речь идет об очень узко очерченной ситуации, то естественно пытаться добиться наибольшей ценности всех произведённых продуктов с учётом ограничений на имеющееся сырьё. Обозначим через x_j переменные управления ($j=1 \dots n$). В данном случае x_j - производимое количество j -го продукта. Тогда целевую функцию, максимум которой мы будем искать, можно задать так:

$$W = \sum_{j=1}^n c_j x_j \quad (\text{суммарная ценность произведённых продуктов}).$$

Перейдём теперь к учёту ограничений. Прежде всего понятно, что производимые количества продуктов не могут быть отрицательными, то есть должны выполняться условия $x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$.

Далее, так как для получения единицы j -го продукта необходимо затратить a_{ij} единиц i -го сырья, то понятно, что для выработки x_j единиц этого продукта потребуется $a_{ij}x_j$ единиц i -го сырья. Поскольку один вид сырья может использоваться для производства различных продуктов, то суммарный расход сырья каждого вида не должен превышать имеющиеся ресурсы, то есть

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i=1 \dots m$$

Окончательно пришли к следующей оптимизационной задаче: найти оптимальное решение, т.е. набор чисел x_j ($j=1 \dots n$), который обеспечит

$$\max W = \max_{\{x_j\}} \sum_{j=1}^n c_j x_j$$

при условиях (ограничениях):

$$\begin{aligned} 1) & x_j \geq 0, & j &= 1 \dots n. \\ 2) & \sum_{j=1}^n a_{ij} x_j \leq b_i, & i &= 1 \dots m. \end{aligned}$$

Всякий набор значений x_1, x_2, \dots, x_n , удовлетворяющий условиям 1 и 2 будем называть допустимым планом (стратегией, управлением или решением). Нас интересует тот допустимый план, который доставляет максимум целевой функции. Будем называть его оптимальным планом (стратегией, управлением, решением). Приведенная задача имеет весьма простую структуру – **целевая функция и ограничения линейны относительно x_j** , то есть задаются функциями простейшего линейного вида. Такая специфика имеет как свои достоинства, так и недостатки. Как установлено, она значительно упрощает процесс математического анализа, гарантирует получение оптимальных решений. Но, с другой стороны, далеко не всегда реальная ситуация хорошо описывается линейными функциями, они могут быть много сложнее по структуре. Тем не менее, класс ситуаций, достаточно хорошо описываемых линейными моделями, весьма широк. Соответствующие оптимизационные задачи получили название задач линейного программирования. Так формулируется типичная задача линейного программирования для экономики, но, как уже отмечалось, исследование операций и, в частности, линейное программирование применяется широко. Рассмотрим пример с решением из области медицины.

Выбор курса лечения

Рассмотрим модель, предложенную Р.Ледли и Л.Лестедом.

Имеются две возможности лечения рака – лучевая терапия и химиотерапия, причем эффективность обоих методов выражена экспертом в некоторых общих единицах. Например, лекарственный препарат обладает эффективностью в 1000 единиц на грамм препарата, а облучение – 1000 единиц в минуту. Допустим, что для выздоровления больного требуется не менее 3000 единиц эффективности. Однако оба метода токсичны. Поэтому ни тот, ни другой нельзя применять неограниченно. Пусть токсичность методов также выражена в общих единицах, например, токсичность лекарства равна 400 единицам на грамм, а токсичность облучения 1000 единицам в минуту. Допустим, что конкретный больной не должен получить в сумме более 2000 таких единиц.

Наконец, известно, что введение одного грамма лекарственного препарата причиняет больному в три раза большие неудобства, чем облучение в течение одной минуты, и, следовательно, если мы ввели x_1 единиц веса лекарств и облучали больного в течение x_2 минут, то причинили ему общее неудобство, равное

$$W = 3x_1 + x_2 \tag{12}$$

Задача состоит в том, чтобы подобрать такое соотношение обоих методов лечения (x_1 и x_2), которое удовлетворяло бы сформулированным выше ограничениям и в то же время причиняло как можно меньше неудобства больному ($\min W$). Такое соотношение назовем оптимальным.

Переходя на математический язык, мы можем сформулировать задачу следующим образом: в плоскости x_1Ox_2 нужно найти такую точку (x_1, x_2) , чтобы величина $W=3x_1+x_2$ была наименьшей, и при этом выполнялись условия:

$$1000x_1+1000x_2 \geq 3000 \quad (13)$$

(ограничение на эффективность) и

$$400x_1+1000x_2 \leq 2000 \quad (14)$$

(ограничение на токсичность).

К этим двум ограничениям необходимо добавить еще одно:

$$x_1 \geq 0 \text{ и } x_2 \geq 0, \quad (15),$$

следующее из того, что x_1 и x_2 по сути задачи не могут принимать отрицательные значения.

Условия (13), (14) и (15) выделяют в плоскости x_1Ox_2 некоторую область, в которой и находится искомая оптимальная точка. Найдём форму этой области. Прежде всего, из условия (15) следует, что искомая точка лежит в первом квадранте. Далее из ограничения (13) следует, что эта точка находится либо на самой прямой

$$x_1 + x_2 = 3 \quad (16),$$

либо выше этой прямой. Аналогично из (14) следует, что точка может находиться либо на прямой

$$2x_1 + 5x_2 = 10 \quad (17),$$

либо ниже этой прямой. Сопоставляя эти условия, получаем, что искомая точка может находиться либо внутри треугольника ABC (рис. 25), либо на его границе.

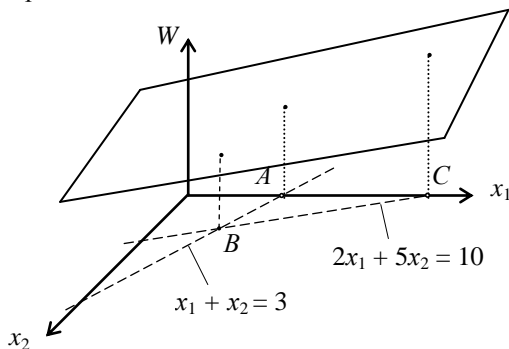


Рис. 25. Подбор оптимального лечения двумя методами (x_1 и x_2)

Итак, из всех возможных точек (x_1, x_2) треугольника ABC (вместе с его границей) нам нужно найти такую, чтобы величина $W = 3x_1 + x_2$ была наименьшей.

Отметим, что уравнение (12) – это уравнение плоскости в пространстве x_1, x_2, W (рис. 25).

Итак, точка (x_1, x_2) пробегает все возможные положения в треугольнике ABC . В теории линейного программирования доказано, что W принимает экстремальное (минимальное или максимальное) значение **на границе треугольника**, точнее в одной вершине этого треугольника или в двух вершинах. В последнем случае любая точка (x_1, x_2) ребра, соединяющего две вершины, также является оптимальным решением.

Таким образом, в задаче подбора курса лечения достаточно найти координаты (x_1, x_2) вершин A, B , и C , подсчитать в этих вершинах значение величины $W = 3x_1 + x_2$, а затем выбрать наименьшее (или равные наименьшие) из этих значений. Координаты точек A и C находятся сразу: $A = (3, 0)$ и $C = (5, 0)$. Найдём координаты точки B . Эта точка лежит на пересечении двух прямых с уравнениями (16) и (17). Следовательно, её координаты должны удовлетворять одновременно обоим уравнениям:

$$\begin{cases} x_1 + x_2 = 3 \\ 2x_1 + 5x_2 = 10 \end{cases}$$

Эта система двух линейных уравнений относительно x_1 и x_2 легко решается, если, например, первое уравнение умножить на два, а затем вычесть его из второго. Мы получим: $x_1 = 5/3, x_2 = 4/3$. Это и есть координаты точки B .

Подсчитаем теперь $W = 3x_1 + x_2$ в точках A, B , и C . Имеем

$$W_A = 9; \quad W_B = 3 \cdot 5/3 + 1 \cdot 4/3 = 19/3 \approx 6,3; \quad W_C = 15.$$

Наименьшее значение целевая функция W (минимум неудобств больному) принимает в точке $B = (5/3, 4/3)$. Следовательно, координаты этой точки и являются искомым решением. Курс лечения будет оптимальным, если ввести $5/3$ грамм лекарственного препарата и провести облечение в течение $4/3$ минуты.

Разумеется в приведённом примере ситуация намеренно упрощена. В реальном случае, например, может быть не один, а несколько лекарственных препаратов. В соответствии с этим может возрасти и число всевозможных ограничений.

Итак, если функция W , наибольшее (или наименьшее) значение которой требуется отыскать, линейна по x_i ($W = a_1x_1 + a_2x_2 + \dots + a_nx_n$) и ограничения записываются также с помощью любых линейных равенств или неравенств, то подобные задачи являются задачами линейного программирования. Когда

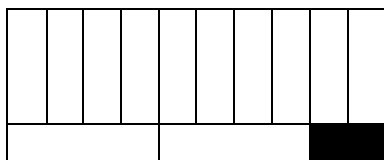
переменных (x_i) и ограничений не 2–3, а больше, то в многомерном пространстве $\{x_i\}$ образуется не треугольник, а многомерная фигура, ограниченная плоскостями. **Оптимальное решение задачи линейного программирования наверняка находится среди «крайних» точек (вершин) этой фигуры.** Число этих точек всегда конечно.

Линейное программирование позволяет решать внешне очень несходные задачи.

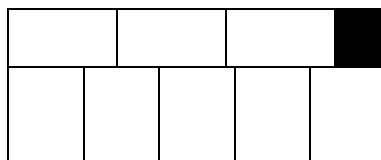
Рациональное размещение

Для организации крупного тепличного хозяйства планируется закупить стандартные теплицы полезной площади 5 x 10 м каждая. В них необходимо разместить не менее 1600 стеллажей типа *A* (размер, включая проходы, 4 x 1 м каждый) и не менее 1600 - типа *B* (2 x 3 м). Необходимо предложить такой план размещения, который позволит выполнить задание при минимальном общем числе закупаемых теплиц.

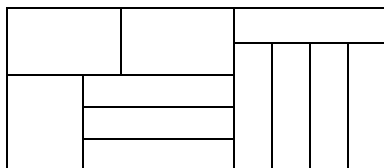
В каждой теплице стеллажи можно разместить по-разному. Например, в теплицу можно поместить лишь один стеллаж типа *A*, а остальную площадь не использовать. Сразу ясно, что такой способ размещения очень плох. Поэтому с самого начала сосредоточим внимание только на «разумных» способах размещения, то есть на таких, где свободной остается минимум площади теплицы, на которой уже нельзя разместить ни одного стеллажа. На рис. 26. приведены такие способы размещения (черным отмечена оставшаяся площадь).



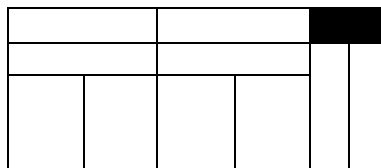
1 способ: 12 стеллажей типа *A*
0 стеллажей типа *B*



2 способ: 0 стеллажей типа *A*
8 стеллажей типа *B*



3 способ: 8 стеллажей типа *A*
3 стеллажа типа *B*



4 способ: 6 стеллажей типа *A*
4 стеллажа типа *B*

Рис. 26. Схемы размещения стеллажей в теплицах

Обозначим через x_i ($i = 1, 2, 3, 4$) количество теплиц, где стеллажи размещены i -м способом. Теперь понятно, что в модели следует оптимизировать – это набор из четырех чисел x_1, x_2, x_3, x_4 . Поскольку мы хотим выполнить план с минимальными затратами на закупку теплиц, целевая функция имеет вид:

$$\min \{x_1 + x_2 + x_3 + x_4\}.$$

Если в одной теплице размещаем стеллажи первым способом, то в нее поместится 12 стеллажей типа A . Если же этот способ применен для x_1 теплиц, то в них поместится $12x_1$ стеллажей типа A . Рассуждая аналогично по отношению к другим способам размещения, можно записать условие выполнения плана по стеллажам типа A :

$$12x_1 + 0x_2 + 8x_3 + 6x_4 \geq 1600$$

и точно так же для размещения стеллажей типа B :

$$0x_1 + 8x_2 + 3x_3 + 4x_4 \geq 1600.$$

В каждом уравнении 4 слагаемых – по числу способов размещения. Кроме того, понятно, что величины x_i ($i = 1 \dots 4$) не должны быть отрицательными.

Окончательная формулировка задачи: найти x_i ($i = 1 \dots 4$), который обеспечит минимум целевой функции

$$W = x_1 + x_2 + x_3 + x_4$$

при условиях:

$$12x_1 + 0x_2 + 8x_3 + 6x_4 \geq 1600$$

$$0x_1 + 8x_2 + 3x_3 + 4x_4 \geq 1600$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0.$$

В результате пришли к задаче линейного программирования. С помощью стандартных приёмов получаем решение – оптимальный план размещения состоит в следующем: 125 теплиц заполняют стеллажами, размещенными вторым способом ($x_2 = 125$), а 200 теплиц – третьим способом ($x_3 = 200$). $x_1 = x_4 = 0$, то есть первый и четвертый способы вообще не используются. Получилось ровно по 1600 стеллажей типа A и B . Конечно далеко не всегда удастся подобрать оптимальное распределение стеллажей почти без лишних площадей, но линейное программирование гарантирует минимум числа покупаемых теплиц. «Вручную» подобный общий план распределения, особенно при более сложной задаче, получить невозможно.

Подобная формализация задачи пригодна, например, для оптимального распределения заданного числа ящиков нескольких размеров в однотипных складских помещениях.

Определение плана перевозок (транспортная задача)

Пусть имеется m предприятий $A_1, A_2, A_3, \dots, A_m$, производящих один и тот же продукт (одного качества) в количествах, равных соответственно $a_1, a_2, a_3, \dots, a_m$. Есть и n потребителей этого продукта, находящихся в пунктах $B_1, B_2, B_3, \dots, B_n$, причем потребности их известны и равны $b_1, b_2, b_3, \dots, b_n$. Предполагается, что суммарный объем потребления равен суммарному объему выпуска продукта на всех предприятиях (задача сбалансирована):

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

Перевозка единицы продукта от i -го предприятия к j -му потребителю ведет к затратам, которые составляют c_{ij} . В этих условиях требуется определить наилучший план перевозок с \min общих затрат.

Построим математическую модель этой ситуации. Через x_{ij} обозначим количество продукта, перевозимого с i -го предприятия к j -му потребителю. Выпишем ограничения, которым должны удовлетворять эти величины. Прежде всего, каждый потребитель должен получить ровно столько продукта, сколько ему требуется, то есть

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

Так как производится столько же, сколько и потребляется, с каждого предприятия продукт должен вывозиться полностью, то есть

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

Понятно также, что перевозимые количества продукта не могут быть отрицательными:

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

В качестве целевой функции, подлежащей минимизации, выступают суммарные затраты на перевозку, определяемые формулой

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

Окончательно приходим к следующей задаче:

Найти x_{ij} , которые обеспечат $\min W$, при выполнении условий:

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

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

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

Пример. Пусть объёмы выпуска предприятий равны следующим величинам: $a_1=145$ т, $a_2=125$ т, $a_3=220$ т, $a_4=135$ т. Объёмы потребления таковы: $b_1=120$ т, $b_2=125$ т, $b_3=130$ т, $b_4=110$ т, $b_5=140$ т. Легко проверить, что задача сбалансирована – объём выпуска равен объёму потребления. Затраты c_{ij} (в руб.) на перевозку единицы продукции (1 тонны) от i -го предприятия к j -му потребителю представлены в таблице.

	$b_1(120)$	$b_2(125)$	$b_3(130)$	$b_4(110)$	$b_5(140)$
$a_1(145)$	18	24	23	27	32
$a_2(125)$	19	20	14	16	26
$a_3(220)$	21	20	17	15	28
$a_4(135)$	15	21	22	19	22

Решение начнём с того, что попробуем подобрать хороший план перевозок, опираясь только на здравый смысл. Будем рассуждать так. Самую дешёвую перевозку, по 14 руб. за 1 т продукта, можно осуществить от второго предприятия к третьему потребителю. Поэтому включим ее в план перевозок с наибольшей возможной интенсивностью, т.е. планируем перевозку 125 т продукта от второго предприятия к третьему потребителю. Следующая минимальная по дороговизне перевозка может быть осуществлена от третьего предприятия к четвертому потребителю 110 т продукта. Рассуждая аналогично, придём к следующему плану перевозки, представленному в таблице.

	$b_1(120)$	$b_2(125)$	$b_3(130)$	$b_4(110)$	$b_5(140)$
$a_1(145)$					140 ↘
$a_2(125)$		5 ↗	125 ↗		
$a_3(220)$		105 ↗	5 ↗	110 ↗	
$a_4(135)$	120 ↗	15 ↗			

Этот план является допустимым, т. к. он позволяет полностью удовлетворить потребности и обеспечивает вывоз продукта с предприятий. Суммарные затраты на его реализацию

составляют: $5 \cdot 24 + 140 \cdot 32 + 125 \cdot 14 + 105 \cdot 20 + 5 \cdot 17 + 110 \cdot 15 + 120 \cdot 15 + 15 \cdot 21 = 12300$ руб.

На практике, к сожалению, нередко наилучший план перевозок отыскивают именно таким способом. Почему «к сожалению», станет ясно из последующего, действительно оптимального плана, полученного методом линейного программирования.

На основании теории линейного программирования, реализованной в пакете стандартных программ для компьютера, получаем решение, представленное в таблице.

	$b_1(120)$	$b_2(125)$	$b_3(130)$	$b_4(110)$	$b_5(140)$
$a_1(145)$	→ 120 ↗	→ 20 ↗			→ 5 ↗
$a_2(125)$			→ 125 ↗		
$a_3(220)$		→ 105 ↗	→ 5 ↗	→ 110 ↗	
$a_4(135)$					→ 135 ↗

Затраты, необходимые для реализации оптимального плана перевозок, составляют: $120 \cdot 18 + 20 \cdot 24 + 5 \cdot 32 + 125 \cdot 14 + 105 \cdot 20 + 5 \cdot 17 + 110 \cdot 15 + 135 \cdot 22 = 11355$ руб. Теперь видно, что по сравнению с первоначальным, казавшимся «хорошим» планом, оптимальный план позволяет сократить затраты более чем на 7%. Причина в том, что в первом плане, начав с максимального использования самых дешёвых путей, мы позже были бы вынуждены остальную продукцию перевозить по слишком дорогим маршрутам.

5.2. Нелинейное программирование

Следующий шаг на пути приближения модели к реальности состоит в том, что целевая функция и (или) ограничения зависят от переменных управления (x_i) нелинейно. Например, производится m видов продукции, причем $f(x_i)$ – затраты не предполагаются пропорциональными x_i – объёму выпуска i -й продукции, т.е. затраты зависят от объема нелинейно.

Другой пример: при увеличении объема выпускаемой продукции спрос на нее может упасть. Цену за единицу продукции, которую при расчетах считали константой, придется понижать, причем цена от объема будет зависеть нелинейно. При формализации подобных ситуаций появятся нелинейные зависимости от переменных управления (x_i).

Оптимизационные задачи, в которых либо целевая функция, либо ограничения, либо и то и другое нелинейны, получили название задач нелинейного программирования. Для них, к

сожалению, нет столь хорошо разработанных методов решения, как в линейном программировании. Поэтому точное решение, обеспечивающее максимум (или минимум) W удастся отыскать далеко не всегда.

Для того чтобы понять, с чем это связано, прежде всего, выясним, на чём может отразиться нелинейность задачи. Предположим, что нелинейны ограничения. Образуется сложная по конфигурации область допустимых решений (на рис. 27 (а) - область U). Сохраняется ли в нелинейных задачах важное свойство линейных задач – конечность числа «крайних» точек (вершин)? Оказывается, нет. Область U имеет бесчисленное множество «крайних» точек.

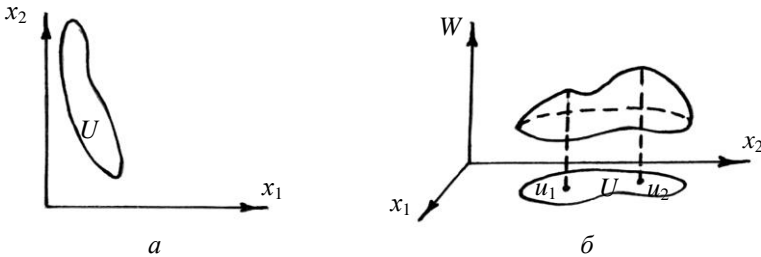


Рис. 27. Область допустимых решений (U) при нелинейных ограничениях (а). Нелинейность целевой функции (б)

Не меньше неприятностей доставляет и нелинейность по x_i целевой функции. Именно из-за нее функция может достигать экстремума (max или min) не в «крайней» точке области U , а где-нибудь внутри. Кроме того, целевая функция может иметь несколько локальных (местных) экстремумов, например, на рис. 27 (б) - в точках u_1 и u_2 (точка x называется точкой локального экстремума, если в ней значение функции больше, чем значение этой функции в достаточно малой окрестности этой точки).

Методы решения задач нелинейного программирования основаны на том, что для любой точки x пространства U мы можем вычислить значение целевой функции W .

1. Наиболее простой метод – «накрывать» область U частой сетью горизонтальных и вертикальных прямых линий по всем переменным x_i . После этого сравнивать значения целевой функции во всех «узловых» точках сети, т.е. посредством перебора найти среди них «узел» x с оптимальным значением целевой функции W . Однако при большом числе переменных (x_i) возникает астрономическое число «узловых» точек.

2. Градиентный метод. Он основан на пошаговом приближении к точке экстремума, двигаясь «шагами» по близким значениям x от меньших значений целевой функции к большим. Этот метод хорошо использовать, когда функция содержит одну вершину. Если же функция имеет более чем одну вершину (исследователь этого не может заранее знать), то дело обстоит намного хуже, так как на этот раз пошаговый поиск может прекратиться не в самой высокой «глобальной», а в какой-нибудь «локальной» вершине W .

В этом случае приходится многократно начинать пошаговый поиск, отправляясь поочерёдно из разных точек области U . Выйдя из какой-нибудь точки, движемся «по градиентам» до точки экстремума. Затем, выбрав, например, случайным образом следующую начальную точку x , вновь ведем поиск экстремума и т. д. Чем больше точек допустимой области U испытано в качестве начальных, тем больше вероятность найти глобальный экстремум. Но гарантии найти его нет: наибольший из найденных экстремумов может быть лишь приближенно принят за значение максимума целевой функции. Приближенно, так как нет и не может быть гарантии, что при таком подходе не окажется пропущенной самая высокая вершина. Проблема в том, что мы «не видим» всей поверхности W в многомерном пространстве, а можем лишь вычислить W для любой конкретной точки x многомерного пространства переменных управления. Просмотреть же все точки невозможно.

На рис.28 и 29 показаны возможности, которые могут представиться при описанном методе поиска экстремума.

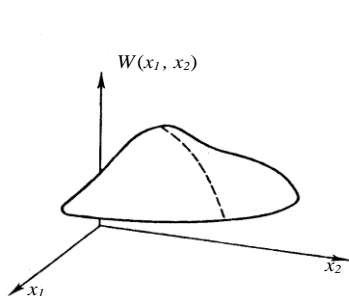


Рис. 28. Целевая функция с одним экстремумом

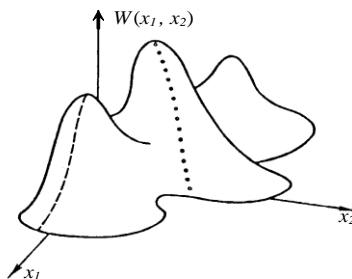


Рис. 29. Целевая функция с несколькими экстремумами

5.3. Динамическое программирование

До сих пор мы рассматривали методику программирования применительно к задачам текущего планирования, когда оптимальный план управления объектом составляется на сравнительно короткий срок и более или менее известны необходимые условия его выполнения.

При перспективном планировании план составляется для нескольких этапов единой операции, т.е. на длительный период, в течение которого возможны существенные изменения, как в условиях производства, так и в соответствующих ресурсах. Причем **эти изменения часто зависят от управления объектом на предыдущих этапах операции**. Чтобы учесть динамику процесса управления, его многоэтапность, в модели приходится учитывать специфические ограничения. Число уравнений, неравенств и, главное, допустимых вариантов решения значительно увеличивается. Подобные задачи решаются методом динамического программирования. Это более молодая отрасль оптимального планирования, чем, например, линейное программирование. **Динамическое программирование специально предназначено для оптимизации многошаговых взаимозависимых операций.**

Пример. Пусть некоторому хозяйству на пять лет выделен кредит в размере Q для развития двух отраслей: растениеводства и животноводства. В начале каждого года часть этих средств распределяется между указанными отраслями. Известна отдача, получаемая от вложения средств в каждую отрасль. При этом отдача отраслей может меняться по годам и, что особенно важно, зависеть от предыдущих вложений. Вопрос заключается в том, чтобы для каждого года определить размер средств, которые следует направить на развитие каждой отрасли, причем, **суммарная прибыль хозяйства, полученная от обеих отраслей за пять лет, должна быть максимальной.**

Сформулированная задача о распределении средств между растениеводством и животноводством оказывается задачей на поиск максимума целевой функции, которая имеет вид

$$W = \sum_{i=1}^5 [f'_{i+1}(x_i) + f''_{i+1}(Q_i - x_i)]$$

Здесь Q_i – часть общего кредита, использованная в i -м году. $f'_{i+1}(x_i)$ – прибыль первой отрасли (растениеводства) в $(i+1)$ -м году при условии, что в предыдущем году в нее вложили x_i средств; $f''_{i+1}(Q_i - x_i)$ – аналогично для второй отрасли

(животноводства). Здесь учтено, что если в первую отрасль вложили x_i средств, то для второй их осталось $Q_i - x_i$, причём

$$\sum_{i=1}^5 Q_i = Q.$$

В реальном случае дело может касаться распределения средств не между двумя отраслями, а среди большего количества. Например, растениеводство можно разбить на «подотрасли»: зерновое хозяйство, овощеводство, кормопроизводство и т. д., животноводство – на молочное, откорм крупного рогатого скота, свиноводство, овцеводство и т. д. В число отраслей, которым выделяются деньги, можно также включить механизацию, мелиорацию, строительство.

Специфика и трудность задач, для решения которых целесообразно прибегать к методам динамического программирования, состоит в том, что **оптимальное управление нужно найти в целом для всей последовательности этапов (лет)**. Сравнительно легко сделать выбор для одного шага, значительно сложнее предусмотреть, как он отразится в долгосрочной перспективе. Соображения ближайшей выгоды порой оборачиваются крупными просчетами. Скажем, мы знаем, что наибольшую прибыль от вложения средств дает животноводство, поэтому можно главную часть средств направить именно в эту отрасль. Но подобное решение может оказаться неправильным, если на дело взглянуть с точки зрения перспективы. Лишая средств растениеводство, мы тем самым подрываем развитие не только данной отрасли, но и затрудняем развитие животноводства, поскольку заведомо ослабляем его кормовую базу, возможно, уже в следующем году.

Предвидеть последствия своих многоэтапных действий – значит, предвидеть будущее. Динамическое программирование как раз и позволяет учитывать те выгоды, которые можно получить не на одном каком-либо этапе, а от всего процесса управления динамическим объектом.

Итак, общее правило планирования многоэтапного процесса состоит в том, что решение на каждом шаге должно приниматься с учетом будущих последствий. Но в реальности часто планирование ведется на один шаг вперед. Дело в том, что предусмотреть, как события станут развиваться в будущем, очень трудно – нужно перебрать огромное число вариантов.

Например, сколько вариантов нужно просчитать для решения вышеизложенной задачи? Предположим для простоты, что общий ресурс средств (Q) разделен по 5-ти годам в заранее известном соотношении. Если в пределах одного года мы примем 10%-ю ступеньку деления средств (10% вложений – растениеводство, 90% - животноводство; 20% растениеводство, 80% - животноводство и т.д.), то

для составления пятилетнего плана придется рассмотреть сто тысяч вариантов: подсчитать для каждого из них предполагаемую прибыль и выбрать из этого огромного количества вариант, обеспечивающий наибольшую прибыль хозяйства в целом за пятилетие. Если дополнительно оптимизировать распределение средств Q по 5-ти годам, то число вариантов становится астрономическим.

Идея решения задач динамического программирования основана на том, что среди шагов, на которых приходится принимать решение, есть один – последний, когда не требуется многовариантных расчетов. Нужно только учесть выгоду, которую можно получить именно на этом этапе. Если бы нам каким-либо образом удалось оптимально распределить средства между отраслями для первых четырех лет, то спланировать их размещение для пятого года не составляло бы труда. Действительно, осталось бы разделить остатки средств между двумя отраслями так, чтобы прибыль, полученная в последнем году, была максимальной.

Идея динамического программирования и состоит в том, что **процесс решения (поиска оптимального плана) начинается с последнего шага (года)**. Рассматриваются все возможные ситуации (остаток средств), возможные в результате выполнения предпоследнего шага и для каждой ситуации выбирается «условно» наилучший вариант последнего шага. **Оптимально спланировав последний шаг, отступаем к предыдущему и тоже оцениваем его с тех же позиций.** Таким образом, процесс динамического программирования разворачивается в обратном порядке – от последнего шага к первому, от конца планового периода к его началу. Выигрыш в объеме вычислений здесь достигается за счет того, что **вместо решения сложной глобальной проблемы, раз за разом решаются несравнимо более простые задачи последовательной «условной» оптимизации одного шага.**

Обратим внимание на еще одну особенность, отличающую динамическое программирование от линейного. Тот и другой метод получили свое название не случайно. Напомним, что сфера использования линейного программирования предполагает линейность функции цели и ограничений. То есть предполагается пропорциональная зависимость между величинами: например, что два трактора сделают вдвое больше работы, чем один, а в двух килограммах сена содержится в два раза больше питательных веществ, чем в одном, и т.д. Во многих случаях такое допущение вполне приемлемо, но далеко не всегда. Так, двойная доза удобрения может не только не дать двойной прибавки урожая, но и нанести вред растениям и почве; увеличение в два раза средств на развитие производства зачастую не способно привести к

двойному увеличению прибыли и т.д. Когда предположение о пропорциональности результата исходным действиям явно несправедливо, обращаться к линейному программированию неправомерно. Что же касается динамического программирования, то оно применимо и для решения многих «нелинейных» задач.

Оптимизация пути

Динамическое программирование начнем с простого игрового примера, предложенного Е.С.Вентцель. Предположим, что нам нужно соорудить путь, соединяющий два пункта *A* и *B*, из которых второй лежит к северо-востоку от первого. Для простоты допустим, что прокладка пути состоит из ряда шагов, и на каждом шаге мы можем двигаться либо строго на восток, либо строго на север; любой путь из *A* в *B* представляет собой ступенчатую ломаную линию, отрезки которой параллельны одной из координатных осей (рис. 30).

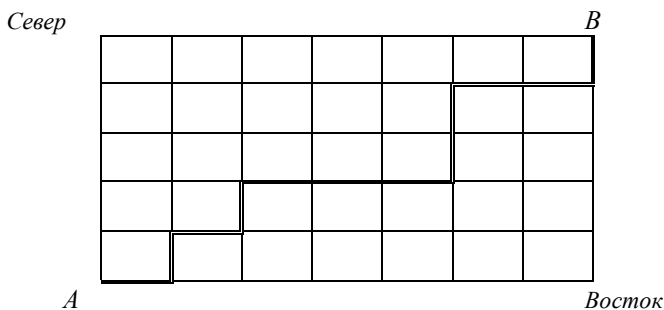


Рис. 30. Путь от пункта *A* в пункт *B*

Затраты на сооружение каждого из таких отрезков заранее известны (они разные). Требуется проложить такой путь из отрезка *A* в *B*, при котором суммарные затраты минимальны.

Как это сделать? Можно поступить одним из двух способов; либо перебрать все возможные варианты пути, и выбрать тот, на котором затраты минимальны (даже при небольшом числе отрезков это очень трудно – слишком много вариантов); либо разделить процесс перехода из *A* в *B* на отдельные шаги (один шаг – один отрезок) и оптимизировать управление по шагам, начиная с последнего. Оказывается, что второй способ гораздо удобнее. Здесь, как и везде в исследовании операций, сказываются преимущества целенаправленного, организованного поиска решения перед «слепым» перебором.

Рассмотрим этот способ решения на примере. Любой путь из A в B состоит из $m=7+5=12$ отрезков, направленных только на восток или на север. Проставим на каждом из отрезков известное число, выражающее стоимость прокладки пути по этому отрезку (рис. 31). Требуется выбрать такой путь из A в B , для которого сумма чисел (затрат), стоящих на всех отрезках пути, минимальна.

Будем рассматривать сооружаемый нами путь как управляемую систему S , перемещающуюся под влиянием управления из начального состояния A в конечное B . Нужно найти оптимальное управление системой. Состояние этой системы перед началом каждого шага будет характеризоваться двумя координатами: восточной (x) и северной (y), обе – целочисленные ($0 \leq x \leq 7$, $0 \leq y \leq 5$). Для каждого из состояний системы (узловой точки прямоугольной сетки) необходимо найти условное оптимальное управление: идти нам из этой точки на север (управление «с») или на восток (управление «в»). Выбирается это управление так, чтобы стоимость всех оставшихся до конца шагов (включая данный) была минимальна.

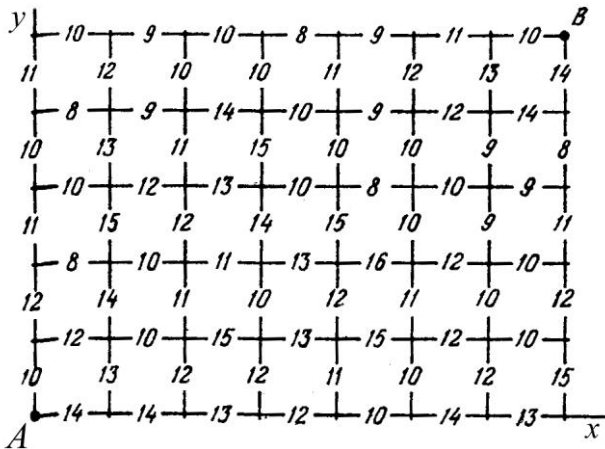
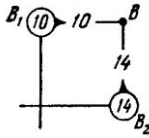


Рис. 31. Стоимости всех отрезков путей от пункта A в пункт B

Процедуру оптимизации будем разворачивать в обратном направлении – от конца к началу. Прежде всего, произведем оптимизацию последнего 12-го шага. Рассмотрим отдельно правый верхний угол нашей прямоугольной сетки (рис. 31), т.е. последний шаг пути:



Где мы можем находиться после 11-го шага? Только там, откуда за 1 (последний) шаг можно попасть в B , то есть в одной из точек B_1 или B_2 . Если мы находимся в точке B_1 , у нас нет выбора (управление вынужденное): надо идти на восток, и это обойдется нам в 10 единиц (условные оптимальные затраты последнего шага). Запишем это число 10 в кружочке у точки B_1 (рисунок), а оптимальное управление покажем короткой стрелкой, исходящей из B_1 и направленной на восток. Для точки B_2 управление тоже вынужденное (север), расход (условные оптимальные затраты) до конца равен 14. Запишем его в кружке у точки B_2 со стрелкой. Таким образом, условная оптимизация последнего шага сделана, и условные оптимальные затраты для каждой из двух возможных точек B_1 и B_2 найдены и записаны в соответствующем кружке.

Теперь оптимизируем предпоследний (11-й) шаг. После предпоследнего (10-го) шага мы могли оказаться в одной из точек C_1 , C_2 , C_3 (рис. 32).

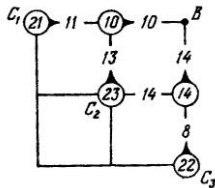


Рис. 32. Схема оптимизации предпоследнего шага пути

Найдем для каждой из них условное оптимальное управление и условные оптимальные затраты. Для точки C_1 управление вынужденное: идти на восток; обойдется это нам до конца пути в 21 единицу (11 в данном шаге, плюс 10, записанных в кружке при B_1). Число 21 записываем в кружке C_1 . Для точки C_2 управление уже не вынужденное: мы можем идти как на восток, так и на север. В первом случае мы затратим на данном шаге 14 единиц и от B_2 до конца – еще 14, всего 28 единиц. Если пойдём на север, то затратим 13+10, всего 23 единицы. Значит, если мы в точке C_2 , то условное оптимальное управление – идти на север (отмечаем это направление стрелкой, а число 23 – условные оптимальные затраты – записываем в кружке у C_2). Для точки C_3

управление снова вынужденное («с»), обойдется оно до конца пути в 22 единицы (ставим стрелку на север, число 22 записываем в кружке у C_3).

Аналогично «пятясь» от предпоследнего шага назад, найдем для каждой точки (всего их $7 \cdot 5 = 35$ не более чем с двумя возможными направлениями движения из точки) условное оптимальное управление («с» или «в»), которое обозначим стрелкой, и условный оптимальный расход до конца пути, который запишем в кружке. Вычисляется он так: расход на данном шаге складывается с уже оптимизированным будущим расходом, записанным в кружке, куда ведет стрелка. Таким образом, на каждом шаге мы оптимизируем только один шаг, а следующие за ним – уже оптимизированы. Конечный результат процедуры оптимизации показан на рис. 33.

Таким образом, условная оптимизация уже выполнена: в какой бы из узловых точек мы ни находились, мы уже знаем, куда идти (стрелка) и во что нам по – минимуму обойдется путь до конца (число в кружке). В том числе, если мы находимся в точке A : в кружке при точке A записан оптимальный расход (цена) на сооружение всего пути из A в B : $W^* = 118$.

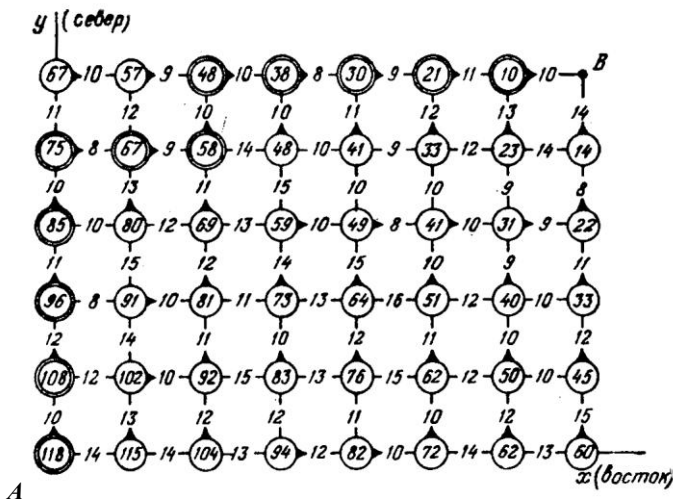


Рис. 33. Конечный результат оптимизации пути

Теперь остается прочитать безусловное оптимальное управление – траекторию, ведущую из A в B самым дешевым способом. Для этого нужно только «идти по стрелкам». Такая оптимальная траектория

отмечена на рисунке дважды обведенными кружками. Соответствующее безусловное оптимальное управление будет:

$$x^*=(c, c, c, c, c, в, в, в, в, в, в, в, в, в, в),$$

то есть первые четыре шага мы должны сделать на север, следующие два на восток, затем опять один на север, и остальные пять на восток. Задача решена.

Заметим, что в ходе условной оптимизации мы можем столкнуться со случаем, когда оба управления для какой-то точки на плоскости являются оптимальными, то есть приводят к одинаковому расходу средств от этой точки до конца. Например, в точке с координатами (5;1) оба управления «с» и «в» являются оптимальными, то есть дают минимальный расход до конца равный 62. Из них мы произвольно выбираем любое (в нашем случае мы выбрали «с»). Такие случаи неоднозначного выбора оптимального управления постоянно встречаются в динамическом программировании. От выбора одного из них, разумеется, может зависеть оптимальное управление всем процессом, но не оптимальный расход средств.

А теперь вернемся к началу и попробуем решить задачу «наивным» способом, выбирая на каждом шаге, начиная с первого, самое выгодное (для этого шага) направление (если таких два – выбираем любое). Таким способом мы получим управление:

$$x=(c, c, в, в, в, в, с, в, в, в, с, в, с, в, с, в).$$

Подсчитаем расходы для этой траектории. Они будут равны $W=10+12+8+10+11+13+15+8+10+9+8+14=128$, что, безусловно, больше, чем $W^*=118$. Причина в том, что «шагнув» в очередной раз по самому дешевому отрезку, мы можем попасть в точку, откуда любой оставшийся путь весьма дорог. В данном случае разница не очень велика, но в других она может быть существенной.

В решенной выше задаче условия были намеренно до крайности упрощены. Разумеется, никто не будет вести железнодорожный путь «по ступенькам», перемещаясь только строго на север или строго на восток. Такое упрощение было сделано для того, чтобы в каждой точке выбирать только из двух управлений «с» или «в». Можно было бы вместо двух возможных направлений ввести их несколько и, кроме того, взять шаги помельче; принципиального значения это не имеет. Разумеется, это усложняет и удлинняет расчеты, но для компьютера подобное усложнение несущественно.

Заметим, что задачи, сходные с рассмотренной выше, очень часто встречаются на практике. Например, при выборе наискорейшего пути между двумя точками или наиболее экономного (в смысле общего расхода горючего) набора, заранее определенных, скорости и высоты для летательного аппарата. Другой пример – необходимость применения

алгоритмов динамического программирования при выполнении т.н. выравнивания нуклеотидных или аминокислотных последовательностей. Эта сложная компьютерная процедура используется в биоинформатике перед оценкой степени гомологичности, родственности генов.

Таким образом, в процессе оптимизации управления методом динамического программирования многошаговый процесс «проходится» дважды: первый раз – от конца к началу, в результате чего находятся условные оптимальные управления и условные оптимальные выигрыши за оставшийся «хвост» процесса; второй раз – от начала к концу, когда нам остается только «прочитать» уже готовые рекомендации и найти безусловное оптимальное управление x^* , состоящее из оптимальных шаговых управлений $x_1^*, x_2^*, \dots, x_m^*$.

Рассмотрим ряд типовых задач, где применим метод динамического программирования и которые «внешне» совершенно не похожи на рассмотренный выше пример.

Задача о распределении ресурсов

В нашем распоряжении имеется ограниченный запас дополнительных средств (ресурсов) K , который должен быть распределен между m популяциями животных P_1, P_2, \dots, P_m . Каждая из популяций P_i при вложении в нее средств (например, дополнительного корма) в размере x приносит дополнительный доход, зависящий от x , то есть $\varphi_i(x)$. Все функции $\varphi_i(x)$ ($i=1..m$) заданы (эти функции неубывающие и, возможно, нелинейные по x). Спрашивается, как нужно распределить запас средств K между популяциями, чтобы в сумме они дали максимальный дополнительный доход?

Эта задача легко решается методом динамического программирования. Хотя в своей постановке она не содержит упоминания о времени, можно все же операцию распределения средств мысленно развернуть в какой-то последовательности, считая за первый шаг вложение средств в популяцию P_1 , за второй – в P_2 и т.д. (хотя их можно поменять местами).

Управляемая система S в данном случае – дополнительные средства (ресурсы), которые обязательно распределяются до конца. Состояние системы S перед каждым «шагом» характеризуется одним числом s – наличным запасом еще не вложенных средств. В этой задаче «шаговыми управлениями» являются средства x_1, x_2, \dots, x_m , выделяемые популяциям. Требуется найти оптимальное управление, то есть такую совокупность чисел x_1, x_2, \dots, x_m ($\sum x_i = K$), при которой суммарный доход максимален:

$$W = \sum_{i=1}^m \varphi_i(x_i) \Rightarrow \max$$

Пример. Исходный запас дополнительных средств $K=10$ (единиц кормов). Требуется его оптимальным образом распределить между пятью популяциями ($m=5$). Для простоты предположим, что вкладываются только целые количества средств. Значения функции дохода $\varphi_i(x)$ (например, в тыс. руб.) приведены в таблице.

В каждом столбце, начиная с какой-то суммы вложений, доходы перестают возрастать (реально это соответствует тому, что каждая популяция способна «потребить» лишь ограниченное количество кормов).

x	$\varphi_1(x)$	$\varphi_2(x)$	$\varphi_3(x)$	$\varphi_4(x)$	$\varphi_5(x)$
0	0	0	0	0	0
1	0,5	0,1	0,6	0,3	1,0
2	1,0	0,5	1,1	0,6	1,2
3	1,4	1,2	1,2	1,3	1,3
4	2,0	1,8	1,4	1,4	1,3
5	2,5	2,5	1,6	1,5	1,3
6	2,8	2,9	1,7	1,5	1,3
7	3,0	3,5	1,8	1,5	1,3
8	3,0	3,5	1,8	1,5	1,3
9	3,0	3,5	1,8	1,5	1,3
10	3,0	3,5	1,8	1,5	1,3

Для получения ответа вначале производят условную оптимизацию так, как это было описано выше, начиная с последнего, 5-го шага.

Оптимальный вариант: надо выделить первой популяции две единицы из десяти, второй – пять единиц, третьей – две, четвертой – ни одной, пятой – одну единицу. Соответствующие суммы доходов отмечены в таблице жирным шрифтом. При этом распределении общий доход будет максимален – 5,6 тыс. руб.

Алгоритмы решения подобных задач динамического программирования несложно реализовать на компьютере.

Задача о загрузке машины

Имеется определенный набор предметов P_1, P_2, \dots, P_n (каждый в единственном экземпляре); известны их веса q_1, q_2, \dots, q_n и стоимости c_1, c_2, \dots, c_n . Грузоподъемность машины равна Q . Спрашивается, какие из предметов нужно взять в машину, чтобы их суммарная стоимость была максимальна при суммарном весе $\leq Q$?

Нетрудно заметить, что эта задача, в сущности, ничем не отличается от предыдущей, но несколько проще ее. Процесс загрузки машины можно представлять себе как состоящий из n шагов; на каждом шаге мы отвечаем на вопрос: брать данный предмет в машину или не брать? Управление на i -м шаге равно единице, если мы данный (i -й) предмет берем, и нулю – если не берем. Значит, на каждом шаге у нас всего два возможных управления.

Характеризовать состояние системы S перед очередным шагом можно весом s , который еще остался в нашем распоряжении до конца (до полной загрузки машины) после того, как предыдущие шаги выполнены (какие-то предметы уже погружены в машину).

Рассмотрим числовой пример. Есть шесть предметов, веса (в тоннах) и стоимости (в тыс. руб.) которых указаны в таблице.

Предмет P_i	P_1	P_2	P_3	P_4	P_5	P_6
Вес q_i	4	7	11	12	16	20
Стоимость c_i	7	10	15	20	27	34

Суммарная грузоподъемность машины $Q=35$ тонн. Требуется указать номера предметов, которые нужно включить в груз, чтобы их суммарная стоимость была максимальна.

Способ решения аналогичен предыдущей задаче, но проще. Оптимальный выигрыш $W^*=57$ тыс. руб. и оптимальные шаговые управления, при которых этот выигрыш достигается: $x_1=0, x_2=1, x_3=0, x_4=1, x_5=1, x_6=0$, то есть загрузить машину надо предметами 2, 4 и 5, суммарный вес которых равен в точности 35 тонн (вообще это не обязательно – при оптимальном выборе грузов может быть и некоторый общий «недогруз»).

5.4. Многокритериальные задачи

Рассмотренные в предыдущих разделах ситуации имели очень важное общее свойство – в каждой из них была единственная целевая функция. Именно единственность этой функции обеспечила возможность создания эффективных методов решения оптимизационных задач. Естественно, возникает вопрос: а хорошо ли такие

оптимизационные модели описывают реальную действительность? Ответ на него неоднозначен.

Да, эти модели могут достаточно хорошо описывать сравнительно простые ситуации, скажем, такие, как обсуждавшиеся выше. Нет, если приходится иметь дело с таким очень часто встречающимся фактом, когда целенаправленная человеческая деятельность преследует сразу несколько целей. В качестве иллюстрации вспомним очень популярный одно время лозунг «Дадим больше товаров лучшего качества по более низкой цене». Этот лозунг в точности характеризует три противоречивые цели, и с этим приходится считаться.

Другой пример – многокритериальный отбор при сравнении линий в сортоиспытании, когда желательно, чтобы отобранные линии имели **одновременно** наибольшую урожайность, процент белка в зерне, самую низкую полегаемость и т.д.

Типичный пример – организация работы предприятия. С одной стороны нам хотелось бы обратить в максимум валовой объем продукции V . Желательно также было бы получить максимальный чистый доход D . Что касается себестоимости S , то ее хотелось бы обратить в минимум, а производительность труда Π – в максимум. При обдумывании задачи может возникнуть еще ряд дополнительных критериев.

Такая множественность показателей эффективности, из которых один желательно обратить в максимум, а другие – в минимум, характерна для любой сколько-нибудь сложной задачи исследования операций. Можно попытаться сформулировать ряд критериев, по которым будет оцениваться фермерское хозяйство, подумать о том, какой из них является главным (теснее всего связанным с целевой направленностью операции), а остальные (дополнительные) расположить в порядке убывающей важности. На этом примере можно убедиться в том, что: а) ни один из показателей не может быть выбран в качестве единственного и б) формулировка системы показателей – не такая уж простая задача. И сами показатели, и их упорядоченность по важности зависят от того, с точки зрения чьих интересов оптимизируется решение.

Итак, типичной для крупномасштабной задачи исследования операций является многокритериальность – наличие ряда количественных показателей W_1, W_2, \dots , одни из которых желательно обратить в максимум, а другие – в минимум.

Спрашивается, можно ли найти решение, одновременно удовлетворяющее всем этим требованиям? Нет. Решение, обращающее в максимум какой-то один показатель, как правило, не обращает ни в

максимум, ни в минимум другие. Поэтому часто применяемая формулировка: «достигнуть максимального эффекта при минимальных затратах» представляет собой не более чем фразу и при научном анализе должна быть отброшена.

Некоторые исследователи пытаются свести **многокритериальную задачу к однокритериальной**: составляют т.н. **индекс** – целевую функцию W , где все показатели W_i являются аргументами, и рассматривают эту функцию как один, «обобщенный» показатель, по которому и оптимизируется решение. Часто такой обобщенный показатель имеет вид дроби, в числителе которой стоят все величины, увеличению которых желательно, а в знаменателе – те, увеличение которых нежелательно. Например, продуктивность и доход – в числителе, время выполнения работы и расходы – в знаменателе и т.д.

Такой способ объединения нескольких показателей в один индекс не может быть однозначно рекомендован, т.к. он предполагает принятие, по крайней мере, одного допущения. А именно, недостаток в одном показателе всегда может быть скомпенсирован за счет другого; например, меньшая продуктивность – за счет более низкой стоимости и т.д. Часто это несправедливо.

Вспомним «критерий для оценки человека», предложенный когда-то Львом Толстым. Он имеет вид дроби, в числителе которой стоят действительные достоинства человека, а в знаменателе – его мнение о себе. С первого взгляда такой подход может оказаться логичным. Но представим себе человека, почти совсем не имеющего достоинств, но совсем не обладающего самомнением. По критерию Л.Н. Толстого такой человек должен иметь бесконечно большую ценность, с чем уж никак нельзя согласиться.

К подобным парадоксальным выводам может привести (и нередко приводит) пользование показателем в виде дроби, где в числителе стоят все величины, увеличение которых желательно, а в знаменателе – те, увеличение которых нежелательно.

Нередко применяется и другой сходный способ составления индекса или «обобщенного показателя эффективности» - он представляет собой «взвешенную сумму» частных показателей, в которую каждый из них W_i входит с каким-то «весом» a_i , отражающим его важность:

$$W = a_1 W_1 + a_2 W_2 + \dots$$

Для тех показателей, которые желательно увеличить, веса a_i берутся положительными, а уменьшить – отрицательными.

При произвольном назначении весов a_1, a_2, \dots этот способ ничем не лучше предыдущего (разве что обобщенный критерий не обращается в бесконечность). Его сторонники ссылаются на то, что и человек,

принимая компромиссное решение, тоже мысленно «взвешивает» все «за» и «против», приписывая больший вес более важным для него факторам. Это, может быть, и так, но, по-видимому, «весовые коэффициенты», с которыми входят в расчет разные показатели, не постоянны, а меняются в зависимости от ситуации.

Рассмотрим это на примере. Человек выходит из дому, чтобы ехать на работу, боится опоздать и размышляет: каким транспортом воспользоваться? Трамвай ходит часто, но идет долго; автобус – быстрее, но с большими интервалами. Можно, конечно взять такси, но это обойдется дорого. Есть еще такое решение: часть пути проехать на метро, а затем взять такси. Но на стоянке может и не быть машин, а добираясь до работы со станции метро пешком, он рискует опоздать больше, чем если бы ехал на автобусе. Как ему поступить?

Перед нами типичная задача исследования операций с двумя критериями (показателями). Первый – среднее время опоздания T , которое хотелось бы сделать минимальным. Второй – стоимость проезда S ; ее тоже желательно сделать минимальной. Но эти два требования, как мы поняли, несовместимы, поэтому человек должен принять компромиссное, приемлемое по обоим критериям решение. В данном случае обобщенный показатель (индекс), который надо обратить в минимум будет выглядеть так:

$$W = a_1 T + a_2 S \rightarrow \min.$$

Но беда в том, что весовые коэффициенты a_1 , a_2 нельзя считать постоянными. Они зависят как от самих величин T и S , так и от обстановки. Например, если человек недавно уже получил выговор за опоздание, коэффициент при T у него, вероятно, увеличится, а на другой день после получки, вероятно, уменьшится коэффициент при S . Если же назначать веса a_1 , a_2 произвольно, то, по существу, столь же произвольным будет и вытекающее из них «оптимальное» решение.

Нельзя надеяться полностью избавиться от субъективности в задачах, связанных с выбором решений. Даже в простейших, однокритериальных задачах она неизбежно присутствует, проявляясь хотя бы в выборе показателя эффективности и математической модели явления. Тем более, неизбежна субъективность при выборе решения в многокритериальной задаче. Правда, бывают редкие случаи, когда достаточно ознакомиться со значениями всех показателей для каждого варианта, чтобы сразу стало ясно, какой из них выбрать. Представим себе, например, что какой-то вариант решения x имеет преимущество над другими по всем показателям; ясно, что именно его следует предпочесть. Но гораздо чаще встречаются случаи, когда с первого

взгляда ситуация неясна: один из показателей «тянет» в одну сторону, другой – в другую.

Другая, практически важная постановка многокритериальных задач – найти оптимальное решение, обеспечивающее максимальное соответствие некому идеалу – приводит к той же проблеме. Например, выбрать среди многих сорт, наиболее близкий к заданному идеалу по комплексу хозяйственно ценных признаков (т.н. модель идеального сорта). У любого сорта есть отклонения от комплексной идеальной модели. Поэтому возникает та же трудно формализуемая задача формирования единой целевой функции: близость каких признаков к их идеальным значениям более, а к каким менее важна и насколько?

Однако, не смотря на это, математический аппарат может помочь без сведения многих критериев к одному. Прежде всего, он позволяет решать «прямые» задачи исследования операций, то есть для любого решения x находить значения показателей эффективности W_1, W_2, \dots, W_k . Для простоты предположим, что все эти величины желательно максимизировать. Пусть в составе множества возможных решений есть два решения x_1 и x_2 такие, что все критерии W_1, W_2, \dots, W_k для первого решения больше или равны соответствующим критериям для второго решения, причем, хотя бы один из них действительно больше. Очевидно, тогда в составе множества решений X нет смысла сохранять решение x_2 : оно «вытесняется» решением x_1 . Выбросим решение x_2 как неконкурентоспособное и перейдем к сравнению других пар решений по всем критериям. В результате такой процедуры отбрасывания заведомо непригодных, невыгодных (по сравнению хотя бы с одним из остальных) решений множество осмысленных решений обычно сильно уменьшается: в нем сохраняются только так называемые **эффективные (иначе «паретовские») решения**, характеризующиеся тем, что ни для одного из них не существует доминирующего (безусловно лучшего) решения среди остальных в этом множестве.

Проиллюстрируем прием выделения паретовских решений на примере задачи с двумя критериями: W_1 и W_2 (оба требуется максимизировать). Множество X состоит из конечного числа n возможных решений x_1, x_2, \dots, x_{20} . Каждому решению соответствуют определенные значения показателей W_1, W_2 ; будем изображать решение точкой на плоскости с координатами W_1, W_2 и занумеруем точки соответственно номеру решения (рис. 35).

Очевидно, из всего исходного множества X эффективными будут только решения x_2, x_5, x_{10}, x_{11} , лежащие на правой верхней границе области возможных решений (точки, соединенные пунктиром). Для всякого другого решения существует хотя бы одно (из этих пяти) доминирующее, для которого оба: W_1 , и W_2 больше, чем для данного

решения. И только для решений, лежащих на правой верхней границе, доминирующих не существует.

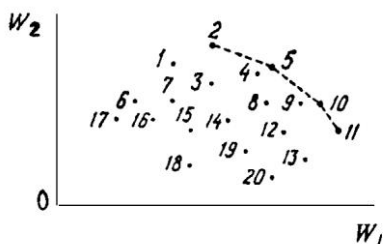


Рис. 35. Распределение решений на плоскости критериев W_1 и W_2

Когда из множества возможных решений выделены эффективные, «переговоры» могут вестись в пределах этого эффективного множества. На рис. 35 его образуют четыре варианта решения: x_2 , x_5 , x_{10} и x_{11} . Ситуация резко упростилась: необходимо сравнить всего четыре варианта. Что касается окончательного выбора, то он всегда остается прерогативой человека. Только человек, с его непревзойденным умением решать неформальные задачи, принимать так называемые «компромиссные решения» (не строго-оптимальные, но приемлемые по ряду критериев) может взять на себя ответственность за окончательный выбор из небольшого числа вариантов.

Аналогично строится множество эффективных решений и в случае, когда показателей не два, а больше (при числе их больше трех, геометрическая интерпретация теряет наглядность).

Существует еще один, часто применяемый способ свести многокритериальную задачу к однокритериальной – это **выделить один (главный) показатель W_1 и стремиться только его обратить в максимум. На все остальные W_2, W_3, \dots предлагается наложить только некоторые ограничения, потребовав, чтобы они были не меньше (или не больше) заранее подобранных w_2, w_3, \dots** Например, при оптимизации плана работы предприятия можно потребовать, чтобы прибыль была максимальна, план по ассортименту – выполнен или немного перевыполнен, а себестоимость продукции – не выше заданной. При таком подходе все показатели, кроме одного – главного, переводятся в разряд заданных условий-ограничений. Известный произвол в назначении границ w_2, w_3, \dots , разумеется, при этом остается; поправки в эти границы тоже могут быть введены в «диалоговом режиме».

Еще один путь построения компромиссного решения можно назвать «методом последовательных уступок». Предположим, что

показатели (например, хозяйственно ценные признаки сравниваемых сортов) W_1, W_2, \dots удалось расположить в порядке убывающей важности. Сначала ищется решение (сорт), обращающее в максимум первый, важнейший показатель, например, урожайность $W_1 = W_1^*$. Затем назначается, исходя из практических соображений, некоторая «уступка» по урожайности ΔW_1 , которую мы согласны сделать для того, чтобы максимизировать второй показатель, например, содержание белка в зерне W_2 . То есть, налагаем на показатель W_1 ограничение: потребуем, чтобы он был не меньше, чем $W_1^* - \Delta W_1$. В этот интервал урожайности попадают несколько сортов. Только среди них ищем решение (сорт), обращающее в максимум W_2 . Далее снова назначим «уступку» в W_2 , ценой которой можно максимизировать W_3 и т.д. Если на каком-то этапе остается всего один сорт, отвечающий очередному ограничению, то, в принципе, можно вернуться к началу алгоритма и выбрать более широкие диапазоны ΔW_i уступок.

Такой способ построения компромиссного решения хорош тем, что здесь сразу видно, ценой какой «уступки» в одном показателе приобретается выигрыш в другом и какова величина этого выигрыша. Недостатком, естественно, является субъективность выбора уступок ΔW_i .

5.5. Проблема оптимизации в условиях неопределенности

В предыдущих разделах мы рассмотрели задачи исследования операций в детерминированном случае, когда показатель эффективности W зависит только от двух групп факторов: заданных, заранее известных α (например, ограничения на ресурсы) и элементов решения x . Реальные задачи исследования операций чаще всего содержат помимо этих двух групп еще одну – **неизвестные факторы**, которые в совокупности обозначим буквой β . Итак, показатель эффективности W зависит от всех трех групп факторов:

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

Так как величина W зависит от неизвестных факторов β , то даже при известных α и x она уже не может быть вычислена, т.е. остается неопределенной. Задача поиска оптимального решения тоже теряет определенность, поскольку нельзя максимизировать неизвестную величину W . И все-таки нам необходимо сделать эту неизвестную величину по возможности максимальной. Поставим перед собой следующую задачу. При заданных условиях α , с учетом неизвестных факторов β , найти такое решение x , которое, по возможности, обеспечивает максимальное значение показателя эффективности W .

Это уже другая задача. Наличие неопределенных факторов переводит ее в новое качество: она превращается в задачу о выборе решения в условиях неопределенности.

Задачи принятия решения в условиях неопределенности встречаются очень часто. Например, 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 точно ни будущий противник, ни вооружение, которым он будет располагать. А решение принимать надо.

Порассуждаем немного о возникшей задаче. Прежде всего, неопределенность есть неопределенность, и ничего хорошего в ней нет. Если условия операции неизвестны, мы не можем также успешно оптимизировать решение, как мы это сделали бы, если бы располагали большей информацией. Поэтому любое решение, принятое в условиях неопределенности, хуже решения, принятого при заранее известных условиях. Однако плохое или хорошее – решение все равно должно быть принято. Наша задача – 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: «Исследование операций представляет собой искусство давать плохие ответы на практические вопросы, на которые даются еще худшие ответы другими методами».

Для того чтобы принимать решения в условиях неопределенности, наука располагает рядом приемов. Какими из них воспользоваться – зависит от того, какова природа неизвестных факторов β , откуда они возникают и как контролируются. Другими словами, с какого вида неопределенностью мы в данной задаче сталкиваемся?

Прежде всего, рассмотрим наиболее благоприятный для исследования, так сказать «хороший» вид неопределенности. Это случай, когда неизвестные **факторы β 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 будем называть стохастическими задачами, а присущую им неопределенность – **стохастической (вероятностной) неопределенностью**.

Рассмотрим более подробно этот «хороший» вид неопределенности. Пусть неизвестные факторы β 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?

Один из способов – **заменить случайные факторы β их средними значениями (математическими ожиданиями)**. Тогда задача становится детерминированной и может быть решена обычными методами. Но весь вопрос в том, насколько случайны эти параметры: если они мало отклоняются от своих математических ожиданий, так поступать можно. Так же обстоит дело и в исследовании операций, где есть задачи, в которых случайностью можно пренебречь. Например, если составляем план снабжения группы предприятий сырьем, можно в первом приближении пренебречь, скажем, случайностью фактической производительности источников сырья (если, разумеется, его производство хорошо отлажено). Тот же приём – пренебречь случайностью и заменить все входящие в задачу случайные величины их математическими ожиданиями – будет уже опрометчивым, если влияние случайности на интересующий нас исход операции существенно.

Рассмотрим пример. Планируется работа ремонтной мастерской, обслуживающей автобазу. Пренебрежем случайностью момента возникновения неисправностей (то есть, заменим случайное время безотказной работы машин его математическим ожиданием) и случайностью времени выполнения ремонта. Скорее всего, такая мастерская, работа которой спланирована без учета случайности, не будет справляться со своей задачей. То есть очень часто встречаются операции, в которые случайность входит по существу, и свести задачу к детерминированной не удастся.

Итак, рассмотрим операцию, где факторы β «существенно случайны» и заметно влияют на показатель эффективности W , в результате чего он тоже «существенно случаен». Попытаемся взять в качестве показателя эффективности среднее значение (математическое ожидание) этой случайной величины $\bar{W} = M[W]$ (например, прибыль хозяйства за конкретный год при среднемноголетней урожайности

высеянных сортов) и выбрать такое решение x , при котором этот усредненный по условиям лет показатель обращается в максимум:

$$\bar{W} = M[W(a, x, \beta)] \rightarrow \max.$$

Заметим, что именно так поступают, выбирая в качестве показателя эффективности в задачах, содержащих неопределенность, не просто «доход», а «средний доход», не просто «время», а «среднее время». Такой подход (называемый **«оптимизацией в среднем»**) иногда вполне оправдан. Действительно, если мы выберем решение так, чтобы среднее значение показателя эффективности обращалось в максимум, то, безусловно, поступим правильнее, чем, если бы выбирали решение наобум.

Что же касается элемента неопределенности, конечно, он сохраняется. Эффективность каждой отдельной операции (за один год), проводимой при конкретных значениях случайных факторов β , может сильно отличаться от ожидаемой как в большую, так, к сожалению, и в меньшую сторону. Однако, оптимизируя операцию «в среднем», мы, в конечном счете, после многих ее повторений выигрываем больше, чем если бы совсем не пользовались расчётом.

Такая «оптимизация в среднем» очень часто применяется на практике в стохастических задачах исследования операций, и пользуются ею, обычно не задумываясь над ее правильностью. Чтобы этот прием был правильным, нужно, чтобы операция обладала свойством повторяемости, и «недобор» показателя эффективности в одном случае компенсировался его «избытком» в другом. Например, если мы предпринимаем длинный ряд однородных операций, с целью получить максимальный доход (средний по многим годам), то доходы от отдельных операций суммируются; «минус» в одном случае покрывается «плюсом» в другом, и все в порядке.

Однако не всегда это допустимо. 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 определенным количеством машин. 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 службы.

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 \bar{T} – среднее время ожидания для больного минимально. Однако 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 \bar{T} , мы рискуем дать предпочтение тому алгоритму, при котором среднее время ожидания мало, но отдельные больные (в удаленных малонаселенных районах) могут ожидать врача очень долго.

Чтобы избежать таких неприятностей, можно заменить показатель эффективности \bar{T} следующим требованием: фактическое время T ожидания врача должно быть не больше какого-то предельного значения t_0 . Поскольку T – величина случайная, нельзя просто потребовать выполнения условия $T \leq t_0$; можно только потребовать, чтобы это условие выполнялось с максимальной вероятностью. Именно вероятность становится целевой функцией ($W = P$), ее надо обратить в максимум.

Если полученный максимум P недостаточно велик, то необходимо увеличить параметр t_0 , что конечно нежелательно. Другой вариант – за счет привлечения дополнительных ресурсов найти новое решение, которое сделает вероятность P достаточно большой. Настолько большой, что событие $T \leq t_0$ будет практически достоверным. Точнее, назначить какое-то значение λ , близкое к единице (например, 0,99 или 0,95), и потребовать, чтобы условие $T \leq t_0$ выполнялось для любого больного с вероятностью, не меньшей, чем λ :

$$P(T \leq t_0) \geq \lambda.$$

Введение такого ограничения означает, что из области возможных решений X (в данной задаче возможных алгоритмов АСУ) исключаются решения, ему не удовлетворяющие. Ограничения подобного типа называются **стохастическими ограничениями**.

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 этой операции, то есть случайно реализовавшееся слишком малое значение показателя эффективности W ; иногда оно может означать попросту катастрофу. Что толку в том, что операция в среднем (за много лет) приносит большой выигрыш, если в данном, единичном случае она может нас разорить? От таких катастрофических результатов можно спастись введением стохастических ограничений. При достаточно большом значении уровня доверия λ можно быть практически уверенным в том, что угрожающее разорение нас не постигнет.

Итак, вкратце рассмотрен случай «хорошей» (стохастической) неопределенности и в общих чертах освещен вопрос об оптимизации

решения в таких задачах. Но стохастическая неопределенность – это почти определенность, если только известны вероятностные характеристики входящих в задачу случайных факторов. Гораздо хуже обстоит дело, когда **неизвестные факторы β не могут быть изучены и описаны статистическими методами**. Это бывает в двух случаях: либо а) распределение вероятностей для параметров β в принципе существует, но к моменту принятия решения не может быть получено, либо б) распределение вероятностей для параметров β вообще не существует.

Пример ситуации типа а): проектируется информационная компьютерная система, 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 система (или аналогичная ей) уже существовала и функционировала достаточно долгое время. Но к моменту создания проекта такой информации нет, а решение принимать надо! Как быть?

Можно применить следующий прием: оставить некоторые элементы решения x свободными, изменяемыми. Затем выбрать для начала какой-то вариант решения, зная заведомо, что он не самый лучший, и пустить систему в эксплуатацию, 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 называют **адаптивными**. Преимущество адаптивных алгоритмов в том, что они не только избавляют нас от предварительного сбора статистики, но и перестраиваются в ответ на изменение обстановки.

Теперь обратимся к самому трудному и неприятному случаю б), когда у неопределенных факторов β вообще не существует вероятностных характеристик; другими словами, когда их нельзя считать «случайными». Напомним, что 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. Это пример «хорошей» стохастической неопределенности. Однако бывает неопределенность и нестохастического вида, которую мы условно назовем «плохой

неопределенностью»: не имеет смысла говорить о «законах распределения» факторов β или других вероятностных характеристиках.

Допустим, например, что планируется некая торгово-промышленная операция, успех которой зависит от того, юбки какой длины β будут носить женщины через два года. Распределение вероятностей для величины β 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 надели юбки, и в каждом из них зарегистрировать величину β , это вряд ли поможет в нашем прогнозе.

Вероятностное распределение величины β попросту не существует, так как не существует массива однородных опытов, где она обладала бы должной устойчивостью. Это случай «плохой неопределенности».

Как же найти решение в подобном случае? Вообще отказываться от применения математических методов решения в данном случае не стоит. Некоторую пользу предварительные расчеты могут принести даже в таких скверных условиях.

Итак, ищем решение x , когда показатель эффективности W содержит «плохую неопределенность» - параметры β , относительно которых никаких сведений мы не имеем, а можем делать лишь предположения. Попробуем все же решить задачу.

Зададим какие-нибудь более или менее правдоподобные значения параметров β . Тогда задача перейдет в категорию детерминированных и может быть решена обычными методами. Однако радоваться рано. Допустим, что затратив много усилий и времени мы это сделали. Будет ли найденное решение хорошим для других условий β ? Как правило, нет. Поэтому ценность его сугубо ограниченная. В данном случае разумно будет выбрать не решение x , оптимальное для каких-то условий β , а некое **компромиссное решение**, которое, не будучи оптимальным (возможно, ни для каких условий) будет все же приемлемым в целом их диапазоне. В настоящее время полноценной научной теории компромисса не существует, хотя некоторые попытки в этом направлении в теории игр и статистических решений делаются.

Обычно окончательный выбор компромиссного решения осуществляется человеком. Опираясь на предварительные расчеты, в ходе которых оценивается большое число W для разных условий β и разных вариантов решения x , он может сравнить сильные и слабые стороны каждого варианта и на этой основе сделать выбор (x).

Подчеркнем еще одну полезную функцию предварительных математических расчетов в задачах с «плохой неопределенностью»: они

помогают заранее отбросить те решения x , которые при любых условиях β уступают другим, то есть оказываются неконкурентоспособными. В ряде случаев это помогает существенно сузить множество решений, 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 «плохой неопределенностью» всегда полезно сталкивать 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 «плохой неопределенности», надо всегда рассчитывать на худшее и принимать то решение, которое дает максимальный эффект в наихудших условиях. Если в таких 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 β зависят от сознательно действующего лица («разумного противника»), отвечающего на любое наше решение наихудшим для нас образом.

В более нейтральных ситуациях принцип «гарантированного 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. Скорее следует подумать о том, откуда можно было бы взять недостающую информацию. Здесь все способы хороши – лишь бы прояснить положение.

Следует упомянуть еще об одном довольно оригинальном методе, не очень «объективном», 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 «плохой неопределенности» (например, в футурологии). Упрощенно, 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 специальными методами наподобие

статистического материала. Результаты обработки, разумеется, сохраняют субъективный характер, но в гораздо меньшей степени, чем если бы мнение высказывал один эксперт.

Подобного рода экспертные оценки для неизвестных условий могут быть применены и при решении задач исследования операций с «плохой неопределенностью». Каждый из экспертов на глаз оценивает степень правдоподобия различных вариантов условий β , приписывая им какие-то субъективные вероятности. Несмотря на субъективный характер оценок каждого эксперта, усредняя оценки целого коллектива, можно получить нечто более объективное и полезное. Таким образом, задача с «плохой неопределенностью» как бы сводится к обычной стохастической задаче.

Теория игр

Итак, рассмотрим **наихудший вид неопределенности**, когда некоторые параметры β , от которых зависит успех операции, неизвестны, и нет никаких данных, позволяющих судить о том, какие их значения более, а какие – менее вероятны. Неопределенными могут быть как внешние (например, природные) «объективные» условия операции, так и «субъективные» - **сознательные действия** противников, соперников или других лиц. Предсказать, как себя поведут эти лица, еще труднее, чем предсказывать в области случайных явлений. Такого рода задачами занимается специальный раздел математики, носящий название «теория игр». Для биологов и специалистов по сельскому хозяйству важны разделы этой теории, в которых предполагается, что человек ведет «**игру с природой**».

Например, если фермер имеет средства лишь на один – два года, то ему нужны рекомендации для получения гарантированного урожая в текущем году, а не в среднем за много лет. Тогда **вынужденная осторожность превращает любую случайность во врага**. Погода из «нормального» стохастического фактора превращается в такого врага.

Общая схема подобных задач такова. Лицо, принимающее решение, имеет возможность сделать выбор из n возможных действий (например, выбрать сорта для посева, агроприемы и т.д.). Определена полезность каждого действия в зависимости от некоторых условий, о которых известно, что одно из них наверняка выполняется (засуха, холод, комфортные условия выращивания и т.д.). А вот какое – не известно. В дальнейшем эти условия мы будем называть состояниями природы. Иногда представляется возможным провести какой – либо эксперимент, который с некоторой вероятностью дает информацию о том, в каком состоянии находится или будет находиться природа (например, попытаться получить прогноз). Проведение такого

эксперимента может, вообще говоря, вести к дополнительным затратам. В этих условиях следует решить, какое действие лучше всего предпринимать.

Конфликтные ситуации. Наиболее простыми из ситуаций, содержащих «плохую» неопределенность, являются так называемые конфликтные ситуации. Так называются ситуации, в которых **сталкиваются интересы двух (или более) сторон, преследующих разные (иногда противоположные) цели, причем выигрыш каждой стороны зависит от того, как себя поведут другие.**

Примеры конфликтных ситуаций многообразны. К ним, безусловно, принадлежит любая ситуация, складывающаяся в ходе боевых действий, ряд ситуаций в области экономики. Столкновение противоречащих друг другу интересов наблюдается также в судопроизводстве, в спорте, межвидовой борьбе. В какой-то мере противоречивыми являются также взаимоотношения различных ступеней иерархии в сложных системах. В некотором смысле «конфликтной» можно считать и ситуацию с несколькими критериями: каждый из них предъявляет к управлению свои требования и, как правило, эти требования противоречивы.

В теории игр разработана математическая теория конфликтных ситуаций. Ее цель – **разработка рекомендаций по оптимальному поведению участников конфликта.**

Каждая непосредственно взятая из практики конфликтная ситуация очень сложна, и ее анализ затруднен наличием привходящих, несущественных факторов. **Чтобы сделать возможным математический анализ конфликта, строится его математическая модель. Такую модель называют игрой.** Игра ведется по определенным **правилам.** Эти правила указывают «права и обязанности» участников, а также исход игры – выигрыш или проигрыш каждого участника в зависимости от сложившейся обстановки.

Человечество издавна пользуется такими формализованными моделями конфликтов – «играми» в буквальном смысле слова (шашки, шахматы, карточные игры и т.п.). Отсюда и название «теории игр», и ее терминология: конфликтные стороны условно называются «игроками», одно осуществление игры – «партией», исход игры – «выигрышем» или «проигрышем». Будем считать, что выигрыши (проигрыши) участников имеют количественное выражение. Если изначально это не так, то всегда можно его приписать, например, в шахматах считать «выигрыш» за единицу, «проигрыш» - за минус единицу, «ничью» - за нуль).

Развитие игры во времени можно представлять как ряд последовательных «ходов» участников (в простейшем варианте – один

ход). **Ходом** называется **выбор игроком одного из предусмотренных правилами игры действий и его осуществление. Ходы бывают личные и случайные.** При личном ходе игрок сознательно выбирает и осуществляет тот или другой вариант действий (пример – любой ход в шахматах). При случайном ходе выбор осуществляется не только волей игрока, но также каким-то механизмом случайного выбора («бросание монеты») с оптимально подобранной вероятностью и т.д.).

Теоретически дело не изменится, если предположить, что весь алгоритм принятия решений выбран игроком заранее. Это будет значить, что игрок выбрал определенную стратегию. Стратегия бывает чистой и смешанной. **Чистая стратегия состоит только из личных ходов, смешанная включает случайные ходы.** Выбор стратегии означает, что игрок может и не участвовать в игре лично, а передать алгоритм выбора своих ходов незаинтересованному лицу. Стратегия, например, может быть задана машине-автомату в виде программы (именно так играют в шахматы компьютер).

Игра называется **игрой с нулевой суммой**, если сумма выигрышей всех игроков равна нулю (то есть каждый игрок выигрывает только за счет других). Самый простой случай – парная игра с нулевой суммой – называется антагонистической (или игрой со строгим соперничеством). Теория антагонистических игр – наиболее развитый раздел теории игр, с четкими рекомендациями. Ниже мы познакомимся с некоторыми ее понятиями и приемами.

Теория игр, как и всякая математическая модель, имеет свои ограничения. Одним из них является предположение о полной («идеальной») разумности противника. В реальном конфликте зачастую оптимальная стратегия состоит в том, чтобы угадать, в чем противник «глуп», и воспользоваться этой глупостью в свою пользу. Схемы теории игр не включают элементов риска, неизбежно сопровождающего разумные решения в реальных конфликтах. В теории игр чаще выявляется наиболее осторожное, «перестраховочное» поведение участников конфликта. Сознавая эти ограничения и, поэтому, не придерживаясь слепо рекомендаций, полученных игровыми методами, можно все же разумно использовать аппарат теории игр как «совещательный» при выборе решения.

Рассмотрим числовой пример игры с нулевой суммой. Предполагается, что результатом игры является плата, которую в соответствии с правилами проигравший платит выигравшему. Ради простоты ограничимся рассмотрением одноходовых игр, в которых участвуют два игрока A и B , причем проигрыш одного, например, B , равен выигрышу другого, то есть A . Для того чтобы полностью определить такую игру, нужно задать таблицу платежей – **платежную**

матрицу. Поясним это на примере. Пусть задана следующая платежная матрица (таблица).

Игроки	<i>B</i>				
		ход 1	ход 2	ход 3	ход 4
<i>A</i>	ход 1	5	4	8	9
	ход 2	0	-2	5	7
	ход 3	1	-1	3	6

Матрица известна обоим игрокам. Игрок *A* должен выбрать одну из строк матрицы (ход). Игрок *B*, не зная результата его выбора, должен выбрать один из столбцов. Число, стоящее на пересечении выбранных ими строки и столбца, определяет выигрыш игрока *A*. Выигрыш игрока *B* равен этому же числу с обратным знаком. Например, если *A* выбрал вторую строку, а *B* – третий столбец, то *A* выигрывает, а *B* проигрывает 5 единиц. Будем считать, что игроки осторожны и целью каждого из них является **максимизация наименьшего возможного (гарантированного) выигрыша.**

Основной вопрос, который возникает в теории игр, состоит в следующем: существует ли наилучший способ игры для каждого из игроков, то есть, имеются ли у них оптимальные стратегии? Сразу видно, что игроку *A* выгоднее всего выбирать ход 1, так как элементы первой строки соответственно больше элементов второй и третьей строк. Точно также игроку *B* выгоднее всего выбирать ход 2, так как элементы второго столбца соответственно меньше элементов остальных столбцов.

В теории игр доказано следующее правило. Если наибольший из минимальных выигрышей для *A* в точности равен наименьшему из возможных максимальных проигрышей для *B*, то есть если минимум в какой-нибудь строке платежной матрицы совпадает с максимумом в соответствующем столбце, то эти строка и столбец являются оптимальными чистыми стратегиями игроков. Точка их пересечения называется **седловой точкой платежной матрицы.** В последнем примере седловой точкой является число 4.

Следовательно, благодаря специфическому свойству данной платежной матрицы – наличию в ней седловой точки, найдены оптимальные чистые стратегии игроков *A* – всегда выбирать ход 1, *B* – ход 2. Число 4 в этом случае носит название цены игры. Смысл этого термина такой: цена игры – это та плата, которую получает оптимально играющий игрок, играя с другим оптимально играющим игроком. Ясно, что ход 1 игрока *A* обеспечивает ему выигрыш не менее 4, а ход 2 игрока *B* гарантирует ему проигрыш не более 4 (игроки *A* и *B* не обязательно равноправны).

Но далеко не каждая платежная матрица имеет седловую точку. Например, матрица

1	-1
-1	1

седловой точки не имеет. Как же находить оптимальные стратегии игрокам, если платежные матрицы не обладают приведенными выше свойствами?

В теории игр доказано, что в отсутствие седловой точки залог успеха (при многократной игре с одной и той же матрицей) состоит в подборе оптимальной смешанной стратегии. То есть для всякой игры с нулевой суммой всегда существуют оптимальные смешанные стратегии. Их подбор – это **подбор частот использования нескольких разных ходов**. Даже если игра проводится один раз, то лучше всего для игрока избрать ход, пользуясь найденными частотами случайного их выбора. Чтобы сделать эти рассуждения до конца понятными, обратимся к задаче.

Рассмотрим пример игры без седловой точки и приведем (без доказательства) ее решение. Игра состоит в следующем: два игрока A и B одновременно и не сговариваясь показывают один, два или три пальца. Выигрыш решает общее количество пальцев: если оно четное, выигрывает A и получает у B сумму, равную этому числу; если нечетное, то наоборот, A платит B сумму, равную этому числу. Как поступать игрокам?

Составим матрицу игры. В одной партии у каждого игрока три возможных хода: показать один, два или три пальца. Матрица 3×3 представлена в таблице. В дополнительном правом столбце приведены минимумы строк, а в дополнительной нижней строке – максимумы столбцов. Седловой точки нет.

Игроки		Игрок B			α
		1 палец	2 пальца	3 пальца	
Игрок A	1 палец	2	-3	4	-3
	2 пальца	-3	4	-5	-5
	3 пальца	4	-5	6	-5
	β	4	4	6	

Нижняя цена игры $\alpha=-3$ и соответствует чистой стратегии A – один палец. Это значит, что при осторожном его поведении мы гарантируем, что он не проиграет больше, чем 3. Положение противника кажется еще хуже: нижняя цена игры $\beta=4$, то есть при осторожном

поведении игрок B проиграет не более 4. В общем, положение не слишком хорошее – ни для той, ни для другой стороны. Нельзя ли его улучшить? Оказывается можно.

Если каждая сторона будет применять не одну какую-то чистую стратегию, а смешанную, в которой первый и третий ходы выбирают с вероятностями $\frac{1}{4}$, а второй – с вероятностью $\frac{1}{2}$, то есть

$$P_A=(1/4, 1/2, 1/4), \quad P_B=(1/4, 1/2, 1/4),$$

то при многократном повторении игры средний выигрыш будет устойчиво равен нулю (значит, игра «справедлива» и одинаково выгодна той и другой стороне). Стратегии P_A, P_B образуют оптимальное решение игры, а ее цена $=0$.

Еще один игровой пример, но уже со схемой решения – задача о встречах.

Саша и Лиза условились встречаться зимой возле кинотеатра. Если Саша придет раньше назначенного времени, то Лизы еще не будет и ему придется мерзнуть. Потери Саши в этом случае можно оценить числом -1 . Если раньше придет Лиза, то ему будет еще хуже: потери равны -4 . В том случае, когда оба приходят одновременно (поздно или рано), потерь нет ни у кого.

Как быть Саше и Лизе? Считая, что перед нами игра двух лиц с нулевой суммой, прежде всего, составим платежную матрицу (таблица):

		Лиза	
		Прийти рано	Прийти поздно
Саша	Прийти рано	0	-1
	Прийти поздно	-4	0

Будем искать оптимальные стратегии участников при многократных встречах. Сначала проверим, нет ли у матрицы седловых точек. Оказывается, что нет. (Минимум в каждой строке отрицателен, а максимумы в столбцах равны 0). Значит, наверняка существуют оптимальные смешанные стратегии для каждого из них.

Пусть Саша выбирает ход «прийти рано» с частотой x , а ход «прийти поздно» – с частотой $1-x$. Аналогично для двух ходов Лизы обозначим частоты ее выбора через y и $1-y$. Средний выигрыш, который получит Саша при многократных свиданиях, составляет:

$$W(x,y) = -4 \cdot y \cdot (1-x) + (-1) \cdot x \cdot (1-y) + 0 \cdot x \cdot y + 0 \cdot (1-x) \cdot (1-y) = 5 \cdot x \cdot y - x - 4 \cdot y.$$

Тогда средний выигрыш Лизы составит: $-W(x,y) = -5xy + x + 4y$. Величину x Саше нужно подобрать так, чтобы выигрыш $W(x,y)$ достиг максимума. Аналогично Лизе – подобрать y , чтобы $-W(x,y)$ был максимален. Вычисляем производную функции W по x и, приравнявая ее

нулю, получаем: $5y-1 = 0$. Производную $-W$ по y также приравняем нулю: $5x-4 = 0$. Отсюда можно найти x и y .

Ответ: $x = 4/5$, $y = 1/5$. Полученный результат объясняется так: Саша должен приходить к кинотеатру в четырех случаях из пяти раньше назначенного времени, то есть каждый раз случайно именно с этими вероятностями принимать решение. Лиза же, наоборот, в четырех случаях из пяти должна опаздывать. Оптимальные смешанные стратегии найдены. Тогда ее средний выигрыш (проигрыш Саши) составит

$$-W = -5 \cdot 4/5 \cdot 1/5 + 4/5 + 4 \cdot 1/5 = 4/5.$$

Любое отклонение от этой смешанной стратегии для Лизы приведет к снижению ее среднего выигрыша (снижение проигрыша для Саши). Аналогичны вредные последствия отклонения от своей смешанной стратегии для Саши.

Игры с природой.

Рассмотрим конкретные примеры этого важного раздела теории игр. Но вначале обсудим критерии успеха в играх вообще и в играх с природой, в частности. Вернемся к одному из трудных вопросов: для данной конкретной ситуации построить отвечающую ей целевую функцию. Решение его выходит за рамки теории игр и относится уже к теории полезности.

Во многих экономических задачах подходящими по смыслу целевыми функциями являются прибыль (или убыток). Наиболее простая цель – это отыскание максимального среднего дохода (или минимального среднего убытка). Предполагаем, что доход зависит от случайно реализовавшегося состояния природы. Тогда средний (по возможным состояниям погоды) доход, точнее математическое ожидание дохода, определяется как сумма произведений величин дохода на вероятности появления тех состояний природы, которые этим доходам соответствуют. Критерием может служить максимум такого математического ожидания.

Критерий этот употребляется далеко не всегда, так как доставляемая им информация слишком усреднена. Например, как уже отмечалось, каждое действие часто оценивается по наихудшему для него состоянию природы. Оптимальным действием считается то, которое приводит к наилучшему (*max*) результату при наихудшем (*min*) состоянии. Такой критерий качества управления носит название максиминного критерия. Ясно, что **максиминная стратегия обеспечивает наилучший ответ на наихудшее состояние природы**, то есть, по сути, это стратегия осторожного, пессимистичного игрока.

Вместо того чтобы рассматривать платежную матрицу при выборе решения в условиях неопределенности, часто используют какую-либо разумно построенную матрицу риска, то есть потерь при разных ходах человека и состояниях природы. Тогда к матрице риска может применяться **минимаксный критерий**, то есть выбирается то действие, которое делает наименьшим максимальный риск. Это тоже осторожная стратегия.

Возможны и другие критерии, учитывающие не наихудшее состояние природы, а ее наилучшее состояние, комбинации наилучшего и наихудшего и т.п. Какой критерий выбрать, зависит от конкретной задачи, а также от человека, который ее решает. Целевая функция зачастую находится в сильной зависимости и от искусства решающего, и от некоторых черт его характера (например, пессимист он или оптимист).

После этих общих рассуждений перейдем к игровой задаче для демонстрации применения конкретной целевой функции, предполагающей одновременно усредненный и минимаксный подходы.

Рассмотрим два возможных варианта погодных условий осенне-зимнего периода, при которых будут храниться корнеплоды столовой свеклы:

P_1 – погода 1-го типа: осень сухая, зима относительно холодная – благоприятные условия для хранения корнеплодов;

P_2 – погода 2-го типа: осень влажная, зима относительно теплая – неблагоприятные условия.

Существует 3 режима хранения корнеплодов в буртах после сбора:

a_1 – режим 1: без использования искусственной пены и без вентиляции;

a_2 – режим 2: без использования пены, но с вентиляцией;

a_3 – режим 3: с использованием пены и с вентиляцией.

Какой-то из трех режимов хранения придется выбрать заранее, не зная будущей погоды. Как найти оптимальное решение?

Прежде всего, составим платежную матрицу, связанную с естественной убылью плодов при хранении (таблица).

Состояние погоды	Режим хранения		
	a_1	a_2	a_3
P_1	0	-1	-3
P_2	-5	-3	-2

Числа в таблице характеризуют риск – потери из-за несоответствия режима хранения погоде. Определить их, конечно, трудно, и сделать это можно разными путями. Потери могут выражаться в некоторых относительных единицах, или в кг, в денежном выражении и т.п. Пусть это ожидаемые потери из-за проявления болезни в результате хранения.

Естественно, что до выбора одного из режимов хранения необходимо получить какие-нибудь сведения о погоде, ожидаемой в осенне-зимнем периоде. Например, прогноз погоды, выраженный одним из нижеследующих ответов:

- p_1 – ожидается холодная зима, без оттепелей;
- p_2 – ожидается умеренно холодная зима с оттепелями;
- p_3 – ожидается теплая, малоснежная зима.

На основании анализа многолетних метеорологических данных известны вероятности соответствия каждого из трех прогнозов, получаемых осенью, реальной реализовавшейся погоде P_1 или P_2 в осенне-зимнем периоде (таблица).

Реальное состояние погоды	Вероятности разных прогнозов (осенью)		
	прогноз p_1	прогноз p_2	прогноз p_3
P_1	0,6	0,25	0,15
P_2	0,2	0,3	0,5

Будем называть стратегией ту совокупность действий (выбор режимов), которую поставим в соответствие трем прогнозам. Например, отметим такие стратегии: $(p_1, p_2, p_3) \rightarrow (a_1, a_1, a_1)$, то есть, какой бы ни был прогноз корнеплоды будут хранить в режиме 1; $(p_1, p_2, p_3) \rightarrow (a_3, a_3, a_3)$, то есть всегда выбирается режим 3; $(p_1, p_2, p_3) \rightarrow (a_1, a_2, a_3)$, то есть, полная вера в прогноз, и т.д.

Легко подсчитать, что всего имеем $3^3=27$ различных стратегий. Какую же из них выбрать? Естественно вычислить средние ожидаемые потери для каждой из 27 стратегий в предположении реализации каждого из двух вариантов погодных условий и сравнить эти потери между собой. В качестве примера оценим R_2 – средние потери при погоде P_2 , если человек придерживается стратегии полной веры в прогноз: $0,2 \cdot (-5) + 0,3 \cdot (-3) + 0,5 \cdot (-2) = -2,9$. То есть средние потери: $R_2 = 2,9$. Для погоды P_1 средние потери при этой стратегии составят $R_1 = 0,7$

Так как любой стратегии сопоставляются два числа – средние потери при каждом из двух возможных вариантов погодных условий, то их легко изобразить геометрически точками, у которых абсциссы R_1 – потери при первом состоянии погоды, а ординаты R_2 – при втором (рис. 36).

Предположим, что человек решил использовать минимаксный критерий для средних потерь, точнее подобрать стратегию, которая обеспечит ему наименьший максимум двух средних потерь: $\min \max (R_1, R_2)$. На рис. 36 видно, что стратегия в этом смысле тем лучше, чем левее и ниже расположена изображающая ее точка. Понятно, что если абсцисса и ордината какой-нибудь точки соответственно меньше, чем абсцисса и ордината другой точки, то последнюю точку (стратегию) можно просто выбросить из дальнейшего рассмотрения.

Применив это рассуждение (паретовский подход – раздел 5.4) установим, что количество точек-стратегий можно существенно уменьшить (рис. 37).

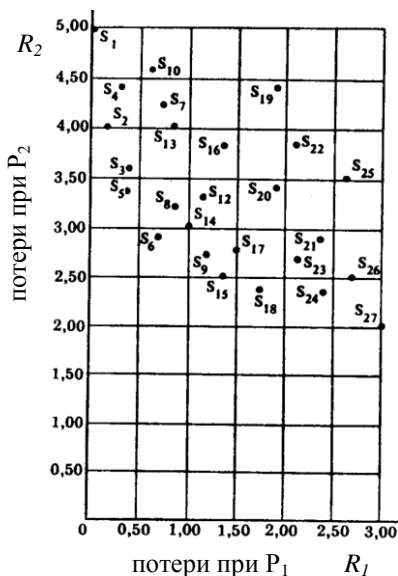


Рис. 36. Распределение средних потерь на плоскости двух состояний природы

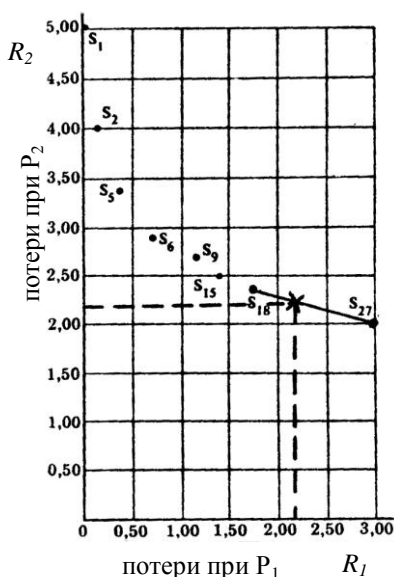


Рис. 37. Оставшиеся стратегии после применения паретовского подхода

Теперь можно привести средние потери при реальных состояниях природы для некоторых оставшихся стратегий S_i (таблица).

Состояние природы (реальная погода)	S_1 a_1, a_1, a_1	S_2 a_1, a_1, a_2	S_5 a_1, a_2, a_2	$\dots S_{27}$ a_3, a_3, a_3
P_1	0	0,15	0,4	$\dots 3$
P_2	5	4	3,4	$\dots 2$
$\max R_1, R_2$	5	4	3,4	$\dots 3$

Покажем теперь, как выбирать стратегии из оставшихся, пользуясь минимаксным критерием.

Точно так же, как и в обычной теории игр, в игре с природой естественно применять стратегии не только в том смысле, как было определено здесь (чистые стратегии), но и смешанные стратегии.

Можно доказать, что смешанная стратегия, изображенная на рис. 37 точкой X обеспечивает минимум максимальных потерь. Чтобы ее реализовать, следует использовать вероятностный механизм, с помощью которого осенью осуществить выбор только между стратегиями S_{18} (a_2, a_3, a_3) и S_{27} (a_3, a_3, a_3). Вероятности выбора S_{18} и S_{27} должны быть обратно пропорциональны расстояниям от точки X до вершин S_{18} и S_{27} , отвечающих этим стратегиям. Такая смешанная стратегия в среднем обеспечит не более, чем 2,2 единицы потерь при любой погоде (P_1 и P_2).

5.6. Заключение: об исследовании операций вообще и в условиях неопределенности в частности.

Задачи, не содержащие неопределенностей, в любой области деятельности человека скорее исключение, чем правило. Адекватное реальности описание проблемы всегда содержит различного типа неопределенности, отражающие то естественное положение, в котором находится исследователь: любое его знание относительно и неточно. Неопределенность проблемы тем выше, чем сильнее зависимость исследуемого объекта от окружающей среды.

Управление системой, функционирующей в условиях неопределенности, требует особой осторожности и обдуманности: выработка наиболее обоснованного комплекса мер важна потому, что в ситуации, когда конечный результат не определен однозначно, на развитие событий можно влиять только принимаемым решением. Принятие неправильного или, по крайней мере, не самого удачного решения всегда связано с потерями, цена которых может быть очень высока. Не случайно идея планомерного совершенствования самих процедур принятия решения зародилась во время второй мировой

войны, когда для выбора стратегии и тактики требовался анализ весьма сложных ситуаций.

Позднее стало очевидно, что общность термина «операция» служит своеобразным отражением другой общности: задачи, возникающие в любой области знания, при всех их качественных различиях, в конечном счете, сводятся к выбору способа действия, варианта плана, параметров конструкций, то есть к принятию решений. И «операция» при этом означает любое целенаправленное действие.

Начиная с 40-х годов проблемам исследования операций посвящается все большее и большее число работ – математических, методологических, а также связанных с анализом конкретных процессов практически во всех сферах научной и производственной деятельности. Первоначально в этих работах господствовал чисто прагматический подход – исследование операций представлялось как собрание различных задач, для которых могли быть использованы однотипные методы решения. Позднее теория исследования операций сложилась в единую научную дисциплину, изучающую определенный класс моделей человеческой деятельности. При решении любой конкретной задачи применение методов этой теории предполагает:

1. Построить математические модели для задач принятия решений и управления в сложных ситуациях или в условиях неопределенности;
2. Изучить на модели взаимосвязи, определяющие возможные последствия принимаемых решений, а также установить критерии эффективности, позволяющие оценивать относительное преимущество того или иного варианта действия.

Модель для исследования операций теснейшим образом связана со спецификой исследуемого процесса, и характер этого процесса вместе с целями моделирования определяет выбор базового математического аппарата. Фиксация этих двух опорных точек – сути процесса и целей математического моделирования – позволяет свести все многообразие ситуаций, требующих того или иного управляющего решения, к вполне ограниченному классу математических постановок. Коснемся существа некоторых из разделов теории исследования операций.

Математическое программирование. В зависимости от вида целевой функции и ограничений на ресурсы используют разные формы программирования (например, линейное, нелинейное, динамическое), которые представляют собой группу вычислительных методов, позволяющих выбрать наилучший план или совокупность действий из множества возможных. Особое распространение получили методы линейного программирования, прежде всего из-за относительной простоты получения оптимальных решений.

Теория игр. В своих прикладных аспектах используется, когда планирование осуществляется в условиях конкуренции, неопределенности или несовершенства знаний о системе и о контактирующей с ней внешней среде. «Предметное» проявление неопределенности представляется как контрплан условного противника (партнера по игре). Примером такой «игры» могут служить взаимоотношения фермера и погоды.

Теория массового обслуживания. Она эффективно используется при исследовании систем, функционированию которых сопутствует процесс образования очередей или задержек обслуживания. Например, организовать работу элеватора таким образом, чтобы при ограниченных ресурсах (обслуживающий персонал, весовое оборудование и пр.) сократить среднее время ожидания разгрузки машин с зерном. При этом прибытие машин на элеватор – случайный процесс.

Сетевой анализ – широко известная группа методов, используемых для планирования крупных разработок и контроля за ходом их выполнения. Например, организовать график строительства крупной фермы таким образом, чтобы выполнение обязательных операций – этапов (подготовка фундамента, подъездных путей, подвоз стройматериалов и т.д.) минимально тормозили друг друга.

Даже из этих предельно сжатых формулировок нетрудно заключить, что многие задачи сельскохозяйственной экономики отлично «вписываются» в методологию исследования операций. Более того, некоторые из них стали хрестоматийными и приводятся в качестве наиболее наглядных примеров в работах общетеоретического (несельскохозяйственного) характера. Это относится к задаче рационального составления комбикорма, иллюстрирующей возможности линейного программирования, к задачам о распределении удобрений, а также об ирригации и складировании, относящимся к компетенции нелинейного и динамического программирования, к проблеме планирования перевозок зерна, снижающего вероятность образования очереди транспортных средств у элеватора (пример классической задачи массового обслуживания).

Контрольные вопросы. 1. Понятие исследования операций, привести примеры задач, перечислить модели и методы, предназначенные для выбора оптимальных решений. 2. Пояснить особенности моделей и привести числовые примеры постановки задач линейного и нелинейного программирования. 3. Пояснить на примерах особенности оптимизационных задач, решаемых методом динамического программирования. 4. Сформулировать задачу динамического программирования на основе модели, описывающей

динамику возрастных групп (раздел 2.6). 5. Каковы сложности решения многокритериальных задач? Привести примеры постановки и методы решения. 6. Пояснить проблему решения оптимизационных задач с учетом влияния неопределенностей различного типа. На примерах пояснить подходы к выбору критериев оптимизации. 7. Привести примеры задач, пояснить смысл критериев и оптимальных стратегий в теории игр.

6. Имитационное моделирование

Все рассмотренные до сих пор модели имели важные общие черты. Для каждой моделируемой ситуации была известна цель (или несколько целей), достижение которой (которых) считалось желательным. Модель обычно представляла собой уравнение или систему уравнений, которая допускала решение – зависимость выходных переменных от входных. Такие модели называют **аналитическими**.

Однако далеко не все ситуации таковы. На современном уровне прикладных исследований часто приходится иметь дело со сложными системами, в которых не только наличествует множество целевых функций, но далеко не все ясно с количественным выражением этих функций. Здесь речь вообще может идти не столько о решении тех или иных оптимизационных задач, сколько об исследовании сложных систем, о прогнозировании их будущих состояний в зависимости от выбираемых стратегий управления.

Коль скоро практика настоятельно потребовала метод для исследования сложных систем, он появился. Этот метод получил название имитационное моделирование (**simulation modeling**).

6.1. Принципы создания имитационной модели

В качестве примера рассмотрим некоторые вопросы, связанные с разработкой имитационной модели Азовского моря. Предварительный анализ, проведенный экспертами, показал, что вся акватория моря может быть разбита на 7 относительно однородных районов. Состояния внутри каждого были описаны с помощью 120 переменных. Среди этих переменных – концентрации химических элементов, биомасса различных видов бактерий, фито-, зоопланктона, основных видов рыб и т.д.

Общая имитационная модель состоит из нескольких блоков, описывающих временную динамику перераспределения между выделенными районами растворенных и взвешенных в воде веществ, а также биогенных элементов, загрязняющих веществ, фито-, зоопланктона, бентоса и рыбы. Для моделирования каждого из блоков используется наиболее подходящий математический аппарат.

Чтобы представить себе сложность общей имитационной модели Азовского моря, достаточно рассмотреть модель любого из блоков, в частности блока, имитирующего пространственно-временную динамику биогенных элементов, то есть соединений азота, фосфора и кремния, определяющих кормовую базу для фито- и зоопланктона. Схематически работа этого блока приведена на рис. 38.

Здесь $y_{ij}(t)$ - содержание i -го биогенного элемента в j -м районе в момент t , $C(t)$ – сток впадающих в море рек, $O(t)$ - количество атмосферных осадков, $A(t)$ - количество смываемого с берегов грунта за счет волн и прилива, $\Phi(t)$ – биомасса фитопланктона, $M_\phi(t)$, $M_3(t)$, $M_p(t)$ – массы отмершего фито-, зоопланктона и рыб соответственно, $T(t)$ – температура воды, $K(t)$ – содержание в воде кислорода, $D(t)$ – характеристика обмена в системе вода – дно.

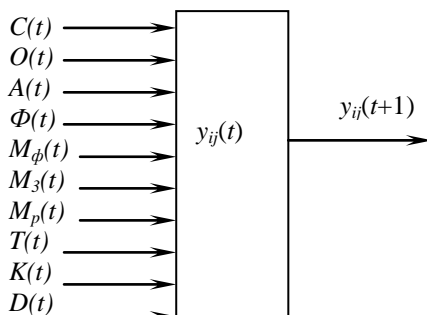


Рис. 38. Схема блока, имитирующего пространственно-временную динамику биогенных элементов Азовского моря

Кроме того, возможны трансформации одних биогенных комплексов в другие, что описывается соответствующими системами дифференциальных уравнений. В частности, для азота эти превращения описываются системой из пяти дифференциальных уравнений первого порядка. В результате блок позволяет получить прогноз – значения y_{ij} в следующий момент ($t+1$) – например, через неделю.

Если учесть, что используется свыше ста переменных, имеется семь районов и рассмотренный блок только один из многих, то можно представить себе исключительную сложность разработанной модели и, кроме того, понять, что аналитическое решение (уравнение) для модели моря получить, естественно, невозможно. При этом следует отметить, что имитационная модель экосистемы Азовского моря в свою очередь – лишь один из блоков имитационной системы водохозяйственного комплекса региона. Имитационная модель Азовского моря используется для проверки возможных последствий тех или иных антропогенных воздействий и долгосрочного прогнозирования.

Само построение таких моделей – сложный процесс. В частности, необходимость знания численных значений коэффициентов, входящих в многочисленные математические уравнения различных блоков, позволила определить те недостающие в настоящий момент

данные, которые должны были быть получены в результате наблюдений, с тем, чтобы выдаваемые прогнозы имели достаточно высокую степень достоверности.

Следует напомнить, что независимо от того, какой метод используется для построения, модели явлений (например, биологических) должны удовлетворять некоторым общим требованиям. Во-первых, результаты (прогнозы состояния моря), получаемые на модели, должны статистически значимо соответствовать экспериментальным данным, а описание, положенное в основу модели, - не быть более сложным, чем это необходимо для получения такого соответствия. Во-вторых, это описание должно содержать информацию о биологическом механизме моделируемого процесса, а модель - обладать возможностями подтверждения результатов тех экспериментов, которые не были использованы при ее построении и подборе коэффициентов в уравнениях (верификация модели).

Итак, суть метода имитационного моделирования состоит в том, что **процесс функционирования сложной системы представляется в виде сложного алгоритма, который реализуется на компьютере**. По результатам реализации с последующей проверкой и исследованием модели могут быть сделаны те или иные выводы относительно исходного процесса.

Перейдем к описанию этапов построения **любой математической модели** системы. Его можно представить себе состоящим из следующих этапов:

1. Формируются основные вопросы о поведении системы, ответы на которые мы хотим получить с помощью модели.

2. Из множества законов, управляющих поведением системы, учитываются те, влияние которых существенно при поиске ответов на поставленные вопросы.

3. В дополнение к этим законам, если необходимо, для моделируемой системы в целом или отдельных ее частей формулируются определенные гипотезы о функционировании. Как правило, эти гипотезы правдоподобны в том смысле, что могут быть приведены некоторые теоретические доводы в пользу их принятия.

4. Гипотезы так же, как и законы, выражаются в форме определенных математических соотношений, которые объединяются в некоторое формальное описание (формулы, алгоритмы).

Критерием адекватности служит практика, которая и определяет, когда может закончиться процесс улучшения модели. Нет надобности говорить, что критерий этот не формализован и в каждом конкретном случае требует специального исследования.

Однако, несмотря на всю привлекательность, описанный подход к построению любых моделей в применении к изучаемым в настоящее время сложным системам обладает определенными недостатками. Прежде всего, определенные трудности могут возникнуть при попытке построить математическую модель очень сложной системы, содержащей много связей между элементами, разнообразные нелинейные ограничения, большое число параметров и т.п. Может статься, что для конкретной сложной системы еще не разработана стройная теория, объясняющая все аспекты ее функционирования, в связи с чем затруднительно формулировать те или иные правдоподобные гипотезы.

Далее, реальные системы зачастую подвержены влиянию различных случайных факторов (погодные условия, случайные ошибки экспериментальных выборок и т.п.). Учет этих факторов аналитическим путем представляет весьма большие, зачастую непреодолимые трудности.

Эти недостатки, систематически возникающие при изучении сложных систем, заставили искать и найти более гибкий метод моделирования – имитационное моделирование. В основе этого метода лежит вполне понятная идея – максимально использовать всю имеющуюся в распоряжении исследователя информацию **об отдельных элементах системы** с тем, чтобы получить возможность обойти аналитические трудности и найти ответ на поставленные вопросы о поведении всей системы.

Если задачи исследования относятся не к выяснению фундаментальных законов и причин, определяющих динамику реальной сложной системы, а к прогнозу и анализу поведения системы, обычно выполняемому в сугубо прикладных целях, то применение имитационного моделирования более чем уместно. Проследим по этапам, как реализуется этот метод с тем, чтобы лучше понять **отличие имитационного моделирования от описанного в предыдущих разделах классического (т.е. аналитического) математического моделирования.**

1. Как и ранее, формируются основные вопросы о поведении сложной системы, ответы на которые мы хотим получить. Множество этих вопросов позволяет задать множество параметров, характеризующих внутреннее состояние сложной системы – вектор состояния. Здесь не всегда помогает даже глубокое знание реальной системы. Так при прогнозировании долгосрочных изменений климата Земли разные группы экспертов предлагали брать за основу несколько отличные векторы состояний атмосферы, почвы, океанов и т.д. В результате были получены несходные выводы по имитационной модели:

в одних случаях следует ожидать потепления, в других – похолодания климата.

2. Осуществляется декомпозиция (разложение) системы на более простые части – блоки. В один блок объединяются «родственные», то есть преобразующиеся по близким правилам, компоненты вектора состояния и процессы, их преобразующие (как в модели Азовского моря). **Декомпозиция завершается, когда процессы внутри каждого блока уже могут быть формализованы в виде обычных аналитических моделей.**

3. Формализуются в аналитической форме физические, химические, биологические и пр. закономерности и «правдоподобные» гипотезы относительно поведения отдельных частей системы. При этом очень важно, что в каждом блоке для их описания может использоваться свой математический аппарат (алгебраические и дифференциальные уравнения, математическую статистику и др.), наиболее удобный для соответствующего блока. Именно блочный принцип дает возможность при построении имитационной модели устанавливать необходимые пропорции между точностью описания каждого блока, обеспеченностью его информацией для настройки и необходимостью достижения общей цели моделирования.

4. Готовятся алгоритмы и компьютерные программы, описывающие функционирование каждого блока. Затем создается единая компьютерная программа, где выходные параметры одних блоков, возможно, «подаются на вход» другим.

Можно сказать, что **под имитационной моделью сложной системы обычно понимают единый комплекс компьютерных программ, описывающий функционирование отдельных блоков системы и правила взаимодействия между ними.**

С учетом целей моделирования выделяются по возможности минимальные наборы «входных» и «выходных» параметров, характеризующих исходные условия функционирования всей системы и прогноз ее реакции на эти условия. Проверка адекватности, верификация модели и сам процесс моделирования поведения всей системы состоит во введении на общий «вход» имитационной модели различных значений этих исходных параметров и последующем анализе значений общих «выходных» результатов – прогнозов.

Использование реализаций случайных величин «внутри» модели (например, требуется учесть случайные погодные условия конкретного региона) делает необходимым применение так называемого **метода Монте-Карло**. Он состоит в многократном проведении компьютерных экспериментов с имитационной моделью: счет («проход») на компьютере с одинаковыми значениями «входных» параметров.

Результаты каждого «прохода», естественно, несколько отличаются, т.к. каждый раз датчик случайных чисел компьютера изменяет «погодные условия». Но характеристики этой случайности (математическое ожидание, дисперсия температуры, количества осадков и пр.) соответствуют разбросу погодных условий конкретного региона для конкретных дат.

Полученные результаты многократного моделирования – «выходные» параметры подвергаются статистическому анализу как данные реальных экспериментов с изучаемой системой.

По существу, каждый «проход» модели состоит в имитации процессов функционирования реальной системы, но в компьютере время ее функционирования сжато в миллионы раз.

6.2. Модели агробиоценоза

Агробиоценоз включает совокупность взаимовлияющих процессов биотического и абиотического характера. Для сельскохозяйственной науки наиболее важная часть агробиоценоза – это посев сельскохозяйственной культуры. При выборе методов моделирования агробиоценоза и степени сложности модели определяющая роль должна отводиться цели моделирования. Например, такой целью может быть прогноз последствий той или иной стратегии проведения сельскохозяйственных мероприятий: орошения, полива, внесения удобрений, подбора сроков посева или посадки растений и пр. с целью получения максимальных урожаев.

Сложность агробиоценоза не позволяет подойти к описанию его функционирования как к процессу, описываемому единым уравнением. Поэтому целесообразно представлять всю систему происходящих в агробиоценозе процессов в виде блочной иерархической структуры. Обычно проводится деление модели на биотический и абиотический блоки. Далее, среди биотических процессов выделяют блок роста и развития посева сельскохозяйственной культуры, блок функционирования почвенной микрофлоры, блок функционирования почвенной фауны, блок развития энтомофауны, блок развития болезней сельскохозяйственных культур, блок взаимодействия сельскохозяйственной культуры с сорняками и др.

Абиотические блоки включают в себя модели, описывающие ряд геофизических процессов, характеристики которых важны для функционирования биотических процессов: формирование теплового, водного режимов почвы и приземных слоев воздуха, концентрации и передвижения биогенных и токсичных солей, различных остатков распада пестицидов, ростовых веществ и метаболитов в почве, концентрация CO_2 в посевах.

Блочная структура моделей дает большие преимущества для моделирования, позволяя изучать, изменять и детализировать одни блоки, не меняя других. Как правило, число параметров, которые входят внутрь блоков, существенно больше числа параметров, которыми блоки соединяются друг с другом. Это один из принципов построения имитационной модели.

Модели продукционного процесса сельскохозяйственных растений, как части агробиоценоза, обычно имеют балансый характер, то есть для каждого вещества проводится расчет всех «притоков» и «оттоков». Например, при расчете водного режима (водный блок) учитываются выпадение осадков (или дождевание), перехват этих осадков надземными органами растений, возможное образование слоя влаги на поверхности почвы, перемещение влаги в почве из одного слоя в другой, обмен с грунтовыми водами, поглощение воды корнями и пр. Таким же образом в модели замыкаются циклы круговорота по углероду, азоту и другим элементам.

На рис. 39 изображена блок-схема модели продуктивности агроэкосистемы, взятая из монографии Н.Ф. Бондаренко и др. «Модели продуктивности экосистем» (1982).

Из блоков, изображенных на рис. 39, наиболее разработаны в настоящее время блоки, описывающие не собственно биологические, а скорее геофизические процессы: влаго- и теплообмен в почве, влаго- и теплоперенос в системе почва – растение – приземный воздух. Это связано, в первую очередь, с большей изученностью геофизических процессов и возможностью их описывать при помощи аппарата дифференциальных уравнений, разработанного для подобных задач в гидро- и аэродинамике. При этом посев формально рассматривается как неоднородная по вертикали пленка, покрывающая поверхность поля.

Подобные имитационные модели можно использовать не только для прогнозирования урожайности и других характеристик, но также для обучения агрономов подбору оперативных воздействий в процессе роста и развития агробиоценоза.

6.3. Модель сои

Эта модель представляет имитационное описание роста развития и формирования урожайности сои и считается наиболее подробной из разработанных за рубежом моделей сельскохозяйственных культур. Для прикладных целей она даже чересчур подробна, однако цели разработки этой модели скорее исследовательские. А именно,

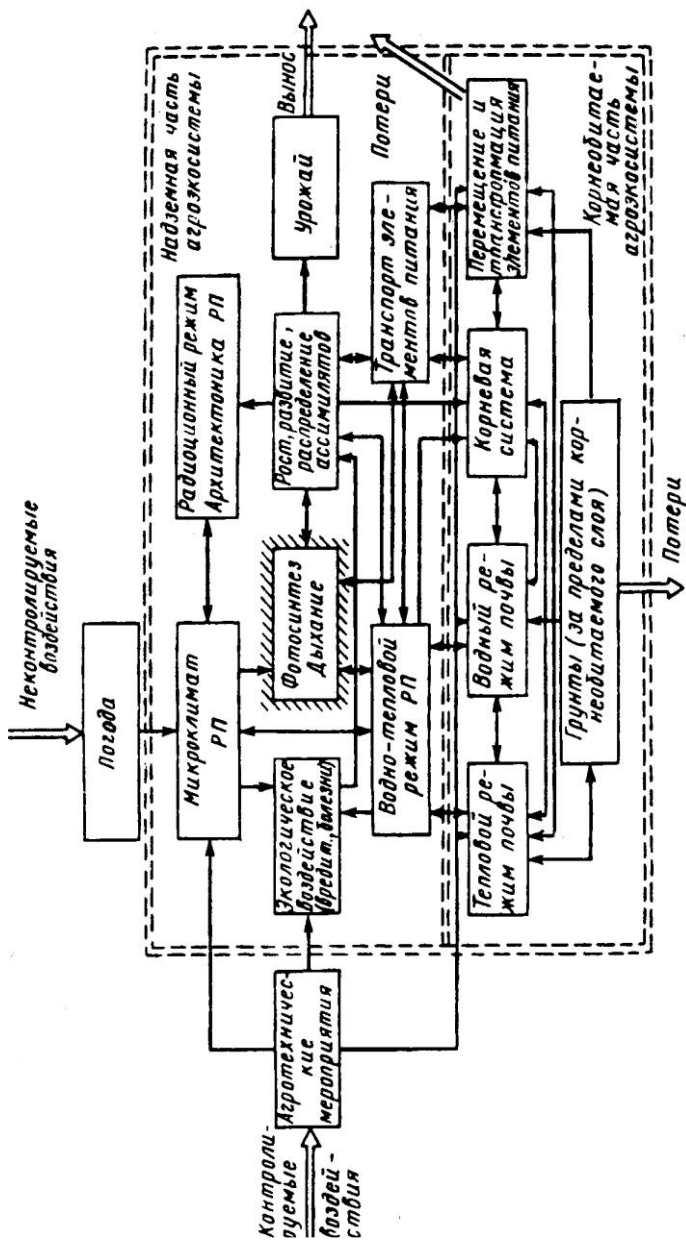


Рис. 39. Блок-схема модели продуктивности агроэкосистемы

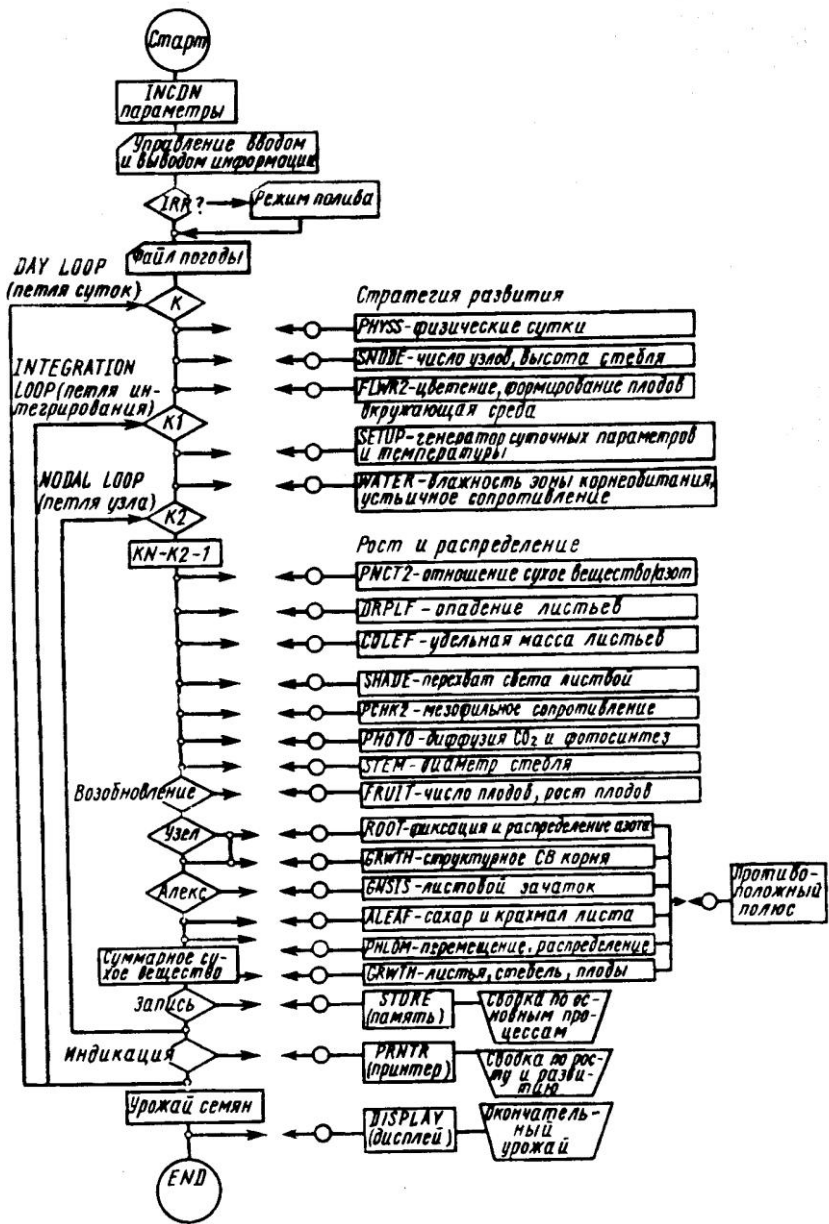


Рис. 40. Блок-схема модели урожайности сои

изучить растение как сложную систему, описать совокупность внутренних процессов и взаимодействий с внешней средой, ответить на вопрос: достаточно ли полны наши знания об этих процессах и насколько они соответствуют реальности.

В модели несколько субмоделей и большое количество входных данных. На рис. 40 изображена упрощенная блок-схема, иллюстрирующая ход вычислительного процесса, взятая из монографии Дж. Франса и Дж. Торнли «Математические модели в сельском хозяйстве» (1987).

В результате сопоставления реальной изменчивости параметров сои, полученной в полевых экспериментах, и предсказанной моделью удалось значительно уточнить, «настроить» имитационную модель этой культуры. Далее возникла возможность использовать модель для исследовательских целей.

Из приведенных примеров ясна степень сложности имитационных моделей для изучения живых систем. В работу по составлению, проверке и использованию одной такой модели вовлечены многие специалисты разных областей: агробиологи, почвоведы, метеорологи, биохимики, экологи, энтомологи и т.д. Математики и программисты, по существу, занимаются обобщением и анализом их рекомендаций. Но они, обычно, не в состоянии понять даже специальную терминологию отдельных областей. Поэтому биологи, специалисты сельского хозяйства, рассчитывающие извлечь пользу из современных методов моделирования, должны быть подготовлены к сотрудничеству с математиками. В частности, понимать принципы и проблемы моделирования.

Несмотря на сравнительную новизну имитационного моделирования как метода исследований сложных систем его результаты иногда существенно влияют не только на принятие научных и хозяйственных решений. Так, в 70-е годы прошлого века были опубликованы результаты глобального моделирования экологических последствий ядерной войны, полученные коллективом ученых под руководством академика Н.Н. Моисеева и повлиявших на политические решения. Модель предсказала неизбежное наступление после войны т.н. «ядерной зимы» с последующей гибелью всего человечества, в том числе победителей и побежденных.

Контрольные вопросы. 1. В чем состоит суть метода имитационного моделирования? 2. Описать области применения и отличия аналитического и имитационного моделирования. 3. Привести этапы построения любой математической модели сложной системы. 4. В чем недостатки метода имитационного моделирования? 5. Как

происходит проверка адекватности построенной модели? 6. Пояснить смысл метода Монте-Карло и его роль в имитационной модели агробиоценоза.

7. Применение непараметрических статистических моделей и методов на примере древесных культур

7.1. Особенности древесных культур как объектов моделирования

При работе с древесными растениями, в частности, плодовыми и декоративными культурами исследователю кроме количественных признаков приходится анализировать большое количество качественных признаков. В таких случаях для анализа результатов наблюдений необходимо применять специальные статистические модели и методы.

Для того чтобы правильно применять те или иные модели необходимо учитывать ряд особенностей объектов исследования:

1) Многолетний образ жизни, включающий длительный ювенильный период, что позволяет оценивать: а) зависимость выражения признаков и их нормы реакции от возраста растения (фактор вариации – возраст); б) влияние года наблюдений (фактор вариации – год);

2) Годичный морфофизиологический цикл развития, включающий период покоя и период вегетации, а также фенофазы и феноинтервалы, что приводит к необходимости оценивать зависимость выражения признаков от календарных сроков наблюдения;

3) Широкая норма реакции по большинству хозяйственно-ценных признаков из-за влияния множества факторов вариации: возраста, фенофазы, почвенно-климатических условий района выращивания, погодных условий текущего и предыдущего года, схемы посадки, подвоя, агротехники (обрезки, полива, минерального питания, системы защиты от вредителей и болезней) и пр.;

4) Множество типов учитываемых показателей и признаков: морфологических (корень, ствол, ветви, побеги, почки, листья, цветки, плоды, семена), хозяйственно-ценных (урожайность, скороплодность, качество плодов и т.п.), устойчивость (к морозам, болезням) и др.;

5) Исследование малых выборок (небольшое количество исходных сортов, гибридные семьи из 10-15 семян и т.п.);

6) Исследование клонов растений: возможность детального анализа модификационной изменчивости.

Статистические модели и методы, используемые при исследовании древесных культур можно подразделить на два типа: одномерные и многомерные.

Одномерные модели (анализ объектов по отдельным признакам): анализ распределения и структуры изменчивости признаков – выяснение достоверности и доли влияния различных факторов, сравнение средних (дисперсионный анализ); анализ сопряженности и

зависимостей между признаками (корреляционный и регрессионный анализ).

Многомерные модели и методы (анализ по множеству признаков): классификация объектов по комплексу признаков (кластерный и таксономический анализ); оценка информативности признаков (метод главных компонент, факторный анализ); прогнозирование выражения признака по косвенным показателям (множественный регрессионный анализ).

Рассмотрим типы шкал, которые используются для описания признаков.

7.2. Шкалы измерений признаков

В основе всякого моделирования лежит оценка переменных (признаков). Чтобы оценить значение переменного, необходимо выбрать шкалу оценки. **Шкала оценки – это способ измерения состояния переменного.** Существует три основных типа шкал оценки признаков: номинальная, порядковая и интервальная.

Эти шкалы отличаются друг от друга по трем основным свойствам:

- 1) наличие или отсутствие правила ранжирования состояний переменного;
- 2) наличие или отсутствие заданного интервала между состояниями переменного;
- 3) набору основных статистических параметров выборки.

Номинальная шкала является низшей шкалой измерения. Номинальные шкалы используют, как правило, для оценки качественных признаков, таких как окраска тех или иных органов, их форма, опушение и т.п.

В широком смысле к качественным относят такие признаки, которые имеют ограниченный ряд состояний, и эти состояния невозможно количественно измерить. Количественными считаются такие переменные, различие между состояниями которых можно измерить. Качественные признаки часто называют номинальными признаками.

Состояние качественного номинального признака называется **модальностью**. В связи с этим, признаки в выборке могут быть мономодальными (отсутствие вариации), бимодальными (две модальности) и полимодальными (три и более модальностей). Например, признак «окраска кожицы плода» в анализируемой выборке имеет 6 модальностей: белая, кремовая, желтая, зеленая, красная, фиолетовая. Поэтому он считается полимодальным.

Исходные данные для анализа номинальных признаков представляют собой наблюдаемые частоты встречаемости модальностей в выборке.

Пример. В коллекции 235 сортов яблоны по признаку «форма плода», наблюдали следующее распределение частот встречаемости модальностей: цилиндрическая – 13 сортов, округлая – 56 сортов, плоскоокруглая – 121 сорт, коническая – 45 сортов.

Единственными математическими связями, уместными по отношению к номинальным шкалам, являются тождество и различие состояний признака у изучаемых объектов. Для характеристики номинальных данных наиболее часто используются пропорция и процентное отношение. Арифметические операции над величинами, измеренными в номинальной шкале, лишены смысла. Единственным показателем средней тенденции является мода (M_o) - модальность, встречающаяся с наибольшей частотой.

Например, в приведенной выше выборке модой является модальность «плоскоокруглая», которая встречается с максимальной частотой 121. Отличительной особенностью моды является то, что число мод в выборке варьирует от одной до нескольких. Если все модальности в выборке встречаются с одинаковой частотой, число мод равно числу модальностей.

Итак, основными статистическими параметрами выборки, при использовании номинальной шкалы являются её объем (N) и показатель средней тенденции, т.е. мода (M_o).

Таким образом, характерными особенностями номинальной шкалы являются:

- а) правило ранжирования модальностей отсутствует (основная особенность);
- б) имеются два статистических параметра выборки: N и M_o ;
- с) интервал между модальностями не определен.

К достоинствам номинальной шкалы можно отнести: а) простоту и быстроту оценки признаков; б) универсальность, то есть, применимость к оценке любого признака, как качественного, так и количественного.

Недостатками номинальной шкалы можно считать: а) субъективность оценки признаков; б) невозможность количественной оценки варьирования признака; в) малая мощность применяемых статистических критериев.

Порядковая (ранговая) шкала, как правило, применима для оценки качественных признаков. Однако в отличие от номинальной шкалы порядковая шкала соответствует таким качественным переменным, для которых характерна упорядоченность, направленность или степень важности их состояний. Например, устойчивость к болезням, выражаемая в баллах или степень опущенности листа: отсутствует, слабое, среднее, сильное. Состояние порядкового признака обычно называют рангом (R_i).

В дополнение к тождеству и различию для порядковых шкал используются связи типа больше или меньше. Однако интервал между рангами не определен и поэтому, как и в случае номинальной шкалы, арифметические операции с рангами не проводят.

В общем виде **рангом R_i наблюдения X_i среди величин X_1, \dots, X_n называют тот порядковый номер, который получит значение X_i при расстановке чисел X_1, \dots, X_n в порядке возрастания или убывания**. Поскольку значения X_1, \dots, X_n в выборке зависят от случая, случайными величинами оказываются и их ранги. В случае равенства X_i для нескольких объектов в выборке, рангом будет среднее арифметическое из соответствующих порядковых номеров этих переменных. Сумма всех рангов в выборке всегда должна быть равна сумме порядковых номеров.

Пример. Исходный вариационный ряд оценок признака у 7 объектов в порядковой шкале (например, степени повреждения штамба плодовых деревьев морозами по 10-ти балльной шкале) – 2; 4; 8; 1; 9; 5; 5.

Ранжированный в порядке возрастания вариационный ряд этих объектов – 1; 2; 4; 5; 5; 8; 9.

Порядковые номера исследованных объектов соответственно – 1; 2; 3; 4; 5; 6; 7.

Ранги объектов - 1; 2; 3; 4,5; 4,5; 6; 7

Сумма рангов: $1+2+3+4,5+4,5+6+7=28$ (сумма порядковых номеров $1+2+3+4+5+6+7=28$).

Переход от самих наблюдений к их рангам сопровождается определенной потерей информации.

Для ранговой шкалы в качестве показателей средней тенденции используют моду и медиану (Me). **Медианой** называется средняя, относительно которой ранжированный ряд распределения делится на две одинаковые половины: по обе стороны от медианы располагается одинаковое число членов ряда. Определить медиану довольно легко. Для этого совокупность наблюдений ранжируют по возрастающим (или по

убывающим) значениям признака, и, если число членов ряда нечетное, то центральная варианта и будет его медианой. При четном числе членов ряда медиана определяется по среднему арифметическому двух соседних вариантов, расположенных в центре ряда. Медиана имеет, по крайней мере, два преимущества перед средним арифметическим: а) она всегда существует в виде точки, разделяющей распределение совокупности пополам (объекты со средним выражением признака могут и не существовать); б) она весьма устойчива к небольшим возмущениям исходного распределения (если имеются выбросы или грубые ошибки их влияние на медиану будет невелико).

Пример. Имеется ранжированный вариационный ряд, содержащий 7 дат – 1, 2, 4, 5, 5, 8, 9. Медианой этого ряда будет центральная варианта под порядковым номером 4, то есть 5.

Для ряда, содержащего 10 дат - 6; 8; 10; 12; 14, 16; 18; 20; 22; 24 – медианой будет полусумма двух его центральных членов, то есть, дат с порядковыми номерами 5 и 6 - $(14+16)/2=15$.

Таким образом, характерными особенностями порядковой шкалы являются:

- а) наличие правила ранжирования состояний переменного (основная особенность);
- б) имеются три основных статистических параметров выборки: N , M_0 и M_e ;
- с) интервал между рангами не определен.

К достоинствам порядковой шкалы обычно относят: а) простоту и быстроту оценки признаков; б) относительную универсальность, то есть, применимость для оценки некоторых качественных и любых количественных признаков.

К недостаткам можно отнести: а) невозможность оценки варьирования признаков; б) малую мощность используемых статистических критериев.

Интервальная шкала оценивает только количественные признаки. В ней возможно отразить, насколько один из объектов отличается от другого по степени выраженности заданного свойства. Отдельное состояние признака в интервальной шкале называется вариантом или датой и, как правило, обозначается x_r .

Для того чтобы задать интервальную шкалу надо определить начальную точку и единицу измерения. Далее при измерении ставят в соответствие каждому объекту число, показывающее, на сколько единиц измерения этот объект отличается от объекта, принятого за начальную

точку (например, температура, в градусах Цельсия или масса в г и т.п.). Количественные шкалы допускают любые арифметические преобразования. В качестве показателей средней тенденции используется M_0 , M_e и среднее арифметическое или другие средние. В качестве показателя варьирования признака обычно используется дисперсия или среднее квадратическое отклонение.

Таким образом, к характерным особенностям интервальной шкалы относят:

- а) наличие правила ранжирования состояний переменного;
- б) имеется набор статистических параметров выборки, основными из которых являются объем выборки, мода, медиана, среднее арифметическое, дисперсия (σ^2) или среднее квадратическое отклонение (σ);
- с) интервал между состояниями переменного определен (основная особенность).

Достоинства интервальной шкалы: а) максимальная мощность используемых статистических критериев; б) объективность оценки переменных.

Недостатки интервальной шкалы: а) относительная сложность оценки переменных; б) невозможность использования для оценки большинства качественных признаков.

В заключение проведем сравнительный анализ различных шкал оценки признаков.

Свойство	Тип шкалы		
	Номинальная	Порядковая	Интервальная
Правило ранжирования состояний переменного	Отсутствует	Имеется	Имеется
Основные статистические параметры выборки	N, M_0	N, M_0, M_e	$N, M_0, M_e, X, \sigma^2, \sigma$
Интервал между состояниями переменного	Не определен	Не определен	Определен

Из таблицы следует, что по совокупности свойств той или иной шкалы они четко разливаются между собой.

7.3. Унификация шкал признаков

При проведении многомерного моделирования, предполагается, что все переменные измерены в одной шкале. Для преобразования исходных данных в единую шкалу используют приемы унификации данных. Существует три пути унификации данных.

1. Сведение всех признаков к номинальной шкале. Этот путь всегда осуществим, но он приводит к большим потерям информации, так как признаки шкал более высокого порядка выражают в шкале признака более низкого порядка. Например, производится сведение всех признаков, вовлекаемых в многомерный анализ, к двоичным переменным: введение вместо каждой исходной случайной переменной серии случайных величины, принимающих только два значения: 0 и 1.

2. Сведение всех признаков к порядковой шкале. Этот путь наиболее часто применяется для унификации данных. Однако он выполнен не для всех переменных и используя его также теряется определенная информация. При этом признаки, измеренные в интервальной шкале, легко перевести в порядковую шкалу. Эту процедуру можно провести, непосредственно заменяя числа на их ранги, или предварительно разбить вариационный ряд на классы и затем заменять числа на порядковые номера соответствующих классов.

Гораздо реже и сложнее можно перевести модальности в ранги. Однако в ряде случаев эта процедура имеет смысл, особенно для качественных признаков, модальности которых можно упорядочить по какому-либо правилу.

Пример. Качественный признак «форма листовой пластинки» сливы имеет 9 модальностей: широкоовальная (1), овальная (2), узкоовальная (3), широкояйцевидная (4), овальнояйцевидная (5), узкоовальнояйцевидная (6), широкообратнояйцевидная (7), овальнообратнояйцевидная (8) и узкоовальнообратнояйцевидная (9).

Однако можно заметить, что форма листа объединяет два разных порядковых признака: 1) степень «овальности листа», то есть сжатости относительно центральной жилки: от широкоовальной до узкоовальной; 2) степень «яйцевидности листа»: от обратнояйцевидной до яйцевидной.

Степень овальности можно количественно оценить путем вычисления индекса овальности (отношения максимальной ширины листовой пластинки к её длине). При этом индекс овальности будет варьировать от 1 (округлый лист), 0,75 (широкоовальный лист), 0,50 (овальный лист), 0,25 (узкоовальный лист) до почти 0 (нитевидный лист). Тогда модальности 1,4 и 7

объединяют широкоовальные листья и имеют балл, например 1; модальности 2,5 и 8 объединяют овальные листья и имеют балл 2; модальности 3,6, и 9 объединяют узкоовальные листья и имеют балл 3.

По аналогии степень яйцевидности можно количественно оценить путем вычисления индекса яйцевидности (отношения расстояния от максимальной ширины листовой пластинки до её основания к её длине). При этом индекс яйцевидности будет варьировать от 1 (обратотреугольный лист), 0,75 (обратнойяйцевидный лист), 0,50 (неяйцевидный лист), 0,25 (яйцевидный лист) до почти 0 (треугольный лист). Тогда по степени «яйцевидности» модальности 4,5 и 6 объединяют яйцевидные листья и имеют балл 1, модальности 1,2 и 3 объединяют листья без яйцевидности и без обратнойяйцевидности и имеют балл 2, модальности 7,8 и 9 объединяют обратнойяйцевидные листья и имеют балл 3.

После перевода модальностей в баллы последние ранжируются и переводятся в ранги.

Таким образом, номинальный признак «форма листовой пластинки» был выражен через два порядковых признака: степень овальности и степень яйцевидности.

3. Сведение всех признаков к интервальной шкале. В данном случае речь идет об оцифровке номинальных и порядковых переменных до уровня количественных признаков. При этом все переменные подтягиваются до уровня количественных путем приписывания их градациям числовых значений. Приписываемые значения иногда называют метками.

Оцифровка качественных и порядковых переменных является сложной и не очень надежной процедурой, как в вычислительном, так и статистическом плане.

7.4. Параметрические и непараметрические методы статистики

Все параметрические методы статистики работают с интервальной шкалой, в отличие от непараметрических методов, ориентированных прежде всего для номинальной и порядковой шкал. Поясним отличия этих методов.

При рассмотрении большинства статистических методов предполагается, что наблюдения, о которых идет речь, выражены в интервальной шкале и являются реализациями случайной величины, распределение которой принадлежит некоторому параметрическому

семейству распределений. Например, случайная величина имеет нормальное, пуассоновское, или другое распределение. То есть, мы предполагаем, что известен закон распределения, например, можем предполагать нормальную $N(\mu, \sigma^2)$ модель, но с неизвестными параметрами μ и σ^2 .

Методы оценивания и проверки гипотез позволяют делать выводы о неизвестных параметрах, при этом ценность заключений зависит от степени адекватности исходного предположения о параметрическом семействе, то есть о форме распределения. Однако существуют случайные величины, которые не подчиняются ни одной из распространенных форм распределения. Следовательно, к ним нельзя применить те математические методы, которые разработаны для параметрических распределений. Поэтому для таких признаков разработаны специальные математические модели, которые получили название непараметрических или свободных от распределения.

Таким образом, можно выделить две группы методов статистики: параметрические и непараметрические. Преимущество параметрических методов состоит в том, что для них существует хорошо разработанный математический аппарат. **Параметрические методы** используют для количественных признаков.

Для анализа номинальных и ранговых переменных используются только **непараметрические методы**, которые **не требуют предварительных предположений относительно вида исходного распределения анализируемых случайных величин**. В этом их достоинство. Но есть и недостаток – снижение чувствительности к различиям объектов. Поясним это.

Напомним, что прежде чем приступить к анализу результатов эксперимента, исследователь выдвигает две взаимоисключающие гипотезы. Одна из них является статистической гипотезой, которую исследователь обычно предполагает отклонить (**нулевая гипотеза H_0** : например, изучаемые сорта не отличаются по урожайности). **Альтернативная гипотеза (H_1)** фактически отрицает нулевую гипотезу (т.е. есть сорта, различающиеся по урожайности).

Выделяют два типа статистических ошибок анализа. **Ошибка первого рода** (ошибка α – типа): отклоняется нулевая гипотеза, которая в действительности верна. **Ошибка второго рода** (ошибка β – типа): принимаем нулевую гипотезу, которая в действительности ложная.

Мощностью статистического критерия называется вероятность того, что будет принято правильное решение (H_1) при ложной нулевой гипотезе. Мощность критерия зависит от объема выборки, уровня значимости, конкретной формулировки нулевой и альтернативной гипотез, надежности экспериментальных данных,

приборов и от самого статистического метода. **При равных условиях параметрические методы более мощные, чем непараметрические.** Но мощность непараметрических методов приближается к параметрическим с увеличением объема выборки.

Каждому типу шкалы соответствует своя статистическая техника. Для номинальных шкал часто используется критерий χ^2 (хи-квадрат). Для порядковых шкал – ранговые статистики. Для интервальных шкал – весь арсенал статистических критериев.

7.5. Алгоритмы и примеры вычисления непараметрических критериев

Номинальная шкала.

Оценка степени сходства между объектами по комплексу признаков. Для решения данной проблемы используют показатель сходства, предложенный Сокалом и Снитом (Sokal, Snith, 1963) и таксономическое отношение Е.С.Смирнова (Смирнов, 1964).

Рассмотрим показатель сходства по Сокалу и Сниту, который предполагает одинаковый вклад всех анализируемых признаков в показатель сходства. Этот показатель определяется как частное от деления числа совпадающих признаков у пары сравниваемых объектов на общее число признаков. Он принимает значения от 0 до 1. Так, если при сравнении двух объектов все признаки совпадают, то показатель сходства равен 1.

Пример. Необходимо определить показатель сходства для 3 сортов по трем признакам: окраске плода, опушению побега и окраске бутона.

№ сорта	Окраска плода	Опушение побега	Окраска бутона
1	Желтая	Есть	Белая
2	Красная	Есть	Розовая
3	Фиолетовая	Нет	Красная

Показатель сходства: между 1 и 2 сортом будет равен $1/3 \approx 0,33$;
между 1 и 3 сортом: $0/3 = 0$;
между 2 и 3 сортом: $0/3 = 0$.

В таксономическом анализе Е.С.Смирнова предполагается, что **вес** модальностей признаков различен в зависимости от частот их встречаемости. Чем реже встречается модальность в выборке, тем её вес больше и наоборот. При этом различают веса по присутствию и по

отсутствию одной и той же модальности. Следовательно, учитываются совпадения не только по присутствию тех или иных модальностей признаков, но и по их отсутствию.

$$T_{ij} = \frac{1}{M} \sum_{k=1}^M w_k, \quad \text{где}$$

T_{ij} – коэффициент сходства между i -м и j -м объектами;

M – общее количество модальностей по всем признакам;

w_k – вес k -ой модальности либо по присутствию ее, либо по отсутствию.

Вес по присутствию k - ой модальности (w_k^+) определяют по формуле: $w_k^+ = \frac{N-n_k}{n_k}$; вес по отсутствию: $w_k^- = \frac{n_k}{N-n_k}$

Всякому несовпадению двух объектов по модальностям приписывается один и тот же вес «- 1».

Здесь N – общее число сравниваемых объектов; n_k – число объектов, у которых данная модальность присутствует.

Пример. Среди 10 сортов, 2 имели опушенную кожицу плодов, 8 неопушенную. Тогда: Вес по присутствию опушения $w_k^+ = (10 - 2) / 2 = 4$. Вес по отсутствию опушения $w_k^- = 2 / (10 - 2) = 0,25$.

Поскольку сорта с опушенной кожицей встречаются более редко (2 из 10) вес по присутствию опушения (4) значительно превосходит вес по его отсутствию (0,25).

Пример: Оценка степени сходства между 5 сортами по 3 признакам (таблица).

№ сорта	Окраска листовой пластинки	Опушение листовой пластинки	Форма листовой пластинки
1	зеленая	Есть	овальная
2	антоциановая	Есть	овальная
3	антоциановая	Нет	яйцевидная
4	зеленая	Есть	яйцевидная
5	пестрая	Есть	обратнойяйцевидная

Для вычислений необходимо определить веса по присутствию и отсутствию определенных модальностей. Обозначим признаки соответственно буквами А, В и С, а их модальности подстрочными цифрами. Присутствие модальности будем обозначать 1, а её отсутствие – 0 (таблица).

Кодировка объектов

№ сорта	Окраска листовой пластинки (А)			Опушение (В)		Форма листовой пластинки (С)		
	Зел.	Антоц.	Пестр.	Есть	Нет	Овал.	Яйцев	Обр. яйц.
	A ₁	A ₂	A ₃	B ₁	B ₂	C ₁	C ₂	C ₃
1	1	0	0	1	0	1	0	0
2	0	1	0	1	0	1	0	0
3	0	1	0	0	1	0	1	0
4	1	0	0	1	0	0	1	0
5	0	0	1	1	0	0	0	1

Определим вес по присутствию модальности «зеленая» признака «окраска листовой пластинки».

$$w_{A_1}^+ = \frac{N - n_{A_1}}{n_{A_1}} = \frac{5 - 2}{2} = 1,5$$

Вес по отсутствию этой модальности равен:

$$w_{A_1}^- = \frac{n_{A_1}}{N - n_{A_1}} = \frac{2}{5 - 2} = 0,7$$

Аналогично определяют веса по присутствию и отсутствию для всех модальностей всех признаков (таблица).

вес	Окраска листовой пластинки			Опушение		Форма листовой пластинки		
	Зел.	Антоц.	Пестр.	Есть	Нет	Овал	Яйц.	Обр. яйц.
w_k^+	1,5	1,5	4,00	0,25	4,00	1,5	1,5	4,00
w_k^-	0,7	0,7	0,25	4,00	0,25	0,7	0,7	0,25

Общее число модальностей $M = 3 + 2 + 3 = 8$

Определим коэффициент сходства между сортами 1 и 2 (см. таблицу – кодировка объектов).

$$T_{1,2} = \frac{1}{8} (-1 - 1 + 0,25 + 0,25 + 0,25 + 1,5 + 0,7 + 0,25) = 0,15$$

Поясним, как было получено выражение в скобках. Производим сравнение двух сортов по всем модальностям. Так при сравнении сорта 1 и 2 по модальности A_1 (зеленая) наблюдается несовпадение (1 у 1-го сорта и 0 у 2-го). Следовательно, записываем «-1», поскольку, всякому

несовпадению двух объектов по модальностям приписывается один и тот же вес «- 1». Далее сравниваем модальности A_2 (антоциановая). Здесь также обнаруживается несовпадение (0 у 1-го сорта и 1 у 2-го). Значит, записываем следующее слагаемое тоже «-1». При сравнении сорта 1 и 2 по модальности A_3 (пестрая) наблюдается совпадение по отсутствию этой модальности (0 у 1-го сорта и 0 у 2-го). Следовательно, записываем вес по отсутствию данной модальности, который равен 0,25. Аналогично определяются все слагаемые выражения в скобках.

Подобным образом вычисляют коэффициенты сходства между всеми парами сортов.

Помимо оценки сходства между всеми парами объектов в таксономическом анализе Е.С.Смирнова вычисляется для каждого объекта так называемый **коэффициент оригинальности**. **Коэффициент оригинальности** представляет собой среднюю сумму весов по присутствию и отсутствию модальностей каждого объекта исследуемой совокупности. Этот коэффициент является мерой оригинальности объекта, то есть, он будет тем больше, чем более редкими модальностями обладает объект. Анализ коэффициентов оригинальности может оказаться очень полезным, например, при оценке той или иной исходной коллекции сортов, линий или гибридов, а именно, позволит отобрать образцы, сочетающие комплекс редких модальностей признаков.

Пример. Определим коэффициент оригинальности для первого сорта, характеризующий этот сорт по наличию редких модальностей.

$$T_{1,1} = \frac{1}{8}(1,5 + 0,7 + 0,25 + 0,25 + 0,25 + 1,5 + 0,7 + 0,25) = 0,68$$

Аналогично вычисляют коэффициенты оригинальности для остальных сортов.

Теперь можно построить матрицу коэффициентов сходства и коэффициентов оригинальности (таблица).

№	1	2	3	4	5
1	0,68	0,15	-0,69	0,15	-0,26
2	0,15	0,68	-0,16	-0,38	-0,26
3	-0,69	-0,16	1,61	-0,16	-0,58
4	0,15	-0,38	-0,16	0,68	-0,26
5	-0,26	-0,26	-0,58	-0,26	1,41

Из полученных данных видно, что сорта 1 и 2 и 1 и 4 наиболее сходны между собой, поскольку у них максимальное значения коэффициента сходства (0,15). Сорта 1 и 3 наиболее сильно отличаются по проанализированным признакам. Их коэффициент сходства самый маленький (-0,69). Из всех сортов наиболее оригинален сорт 3. Его коэффициент оригинальности составляет 1,61, что выше, чем у остальных сортов. Следовательно, этот сорт сочетает больше редких модальностей.

К полученной матрице коэффициентов сходства можно применить иерархический и неиерархический кластерный анализ. В частности, иерархическая кластеризация позволяет последовательно объединять сорта сначала с максимальным коэффициентом сходства, а затем и менее сходные между собой. Результат кластерного анализа представляют в виде дендрограммы, характеризующей группировки объектов. Для полученной матрицы дендрограмма кластерного анализа имеет следующий вид (рис. 41)

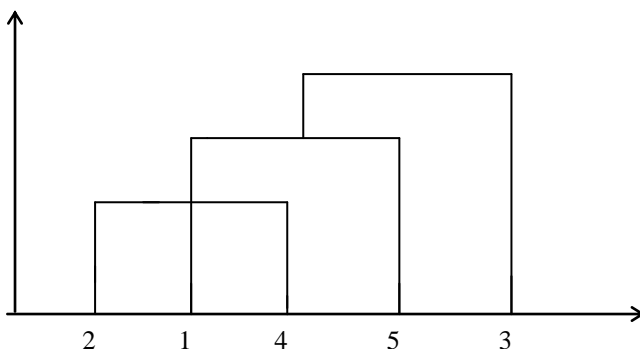


Рис. 41. Дендрограмма сходства 5 объектов

Кластерный анализ позволяет разделить изучаемую выборку объектов на группы, кластеры по степени сходства комплекса признаков. Результаты кластеризации оказываются полезными при решении многих сложных биологических проблем, в частности: 1) классификации таксонов разного ранга: родов, видов, разновидностей, форм, сортов, гибридов, линий, популяций и т.п.; 2) оценки сходства гибридов с родительскими формами; 3) подбора родительских форм для скрещиваний по степени их фенотипического несходства и др.

Кроме того, иногда используется и неиерархическая кластер процедура методом максимального корреляционного пути, который будет рассмотрен ниже.

Ранговая шкала

Критерий Манна-Уитни. Наиболее мощным непараметрическим критерием для оценки различий между центральными параметрами (средними, медианами и т.п.) двух выборок является **U-критерий Манна-Уитни.**

Порядок вычисления этого критерия следующий:

1. Объединение двух групп наблюдений и ранжирование единой выборки. Но, в то же время, для каждого наблюдения одновременно с присвоением ему ранга необходимо помнить (отметить) его принадлежность к исходной группе.

2. Разделение единой выборки на две исходные группы, но уже в ранговой шкале.

3. Определение сумм рангов по каждой выборке.

4. Определение значения расчетного критерия для каждой группы по формулам:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - \sum R_{i(1)} ;$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - \sum R_{i(2)} ;$$

где n_1 – объем первой выборки; n_2 – объем второй выборки; $\sum R_{i(1)}$ – сумма рангов первой выборки; $\sum R_{i(2)}$ – сумма рангов второй выборки.

Несложно показать, что $U_1 = n_1 n_2 - U_2$

5. Если найденные значения критерия (U_1, U_2) входят в интервал для двух пороговых табличных значений U_T , то выборки не различаются по центральному параметрам.

Этот критерий можно использовать и для сравнения двух выборок, имеющих разный объем.

Пример. Необходимо сравнить две группы сеянцев вишни по устойчивости к коккомикозу. В таблице представлена сумма баллов поражения за 5 учетов в течение всей вегетации.

Группа 1	20	18	19	15	14	10	12	17	11	
Группа 2	16	11	9	13	13	11	7	13	9	8

$n_1 = 9$ сеянцев, $n_2 = 10$ сеянцев.

Объединяем обе группы наблюдений и ранжируем в возрастающем порядке:

Суммарный балл	Ранг	Группа	Суммарный балл	Ранг	Группа
7	1	2	13	11	2
8	2	2	13	11	2
9	3,5	2	14	13	1
9	3,5	2	15	14	1
10	5	1	16	15	2
11	7	2	17	16	1
11	7	2	18	17	1
11	7	1	19	18	1
12	9	1	20	19	1
13	11	2			
Сумма				190	

Разделяем группы и определяем суммы рангов (жирный шрифт):

Группа 1	5	7	9	13	14	16	17	18	19		118
Группа 2	1	2	3,5	3,5	7	7	11	11	11	15	72

$$U_1 = 9 \cdot 10 + 9 \cdot (9+1)/2 - 118 = 135 - 118 = 17$$

$$U_2 = 9 \cdot 10 + 10 \cdot (10+1)/2 - 72 = 145 - 72 = 73$$

$$\text{Проверка } U_1 = n_1 n_2 - U_2 = 90 - 73 = 17$$

Обратившись к специальной таблице А для случая $n_1=9$, $n_2=10$ находим, что $20 < U_T < 70$ ($\alpha=0,05$, критерий двухсторонний), следовательно, требуемое для отклонения H_0 значение U должно быть меньше или равно 20 или больше 70. Поскольку расчетные значения U_1 и U_2 равны соответственно 17 и 73 H_0 отклоняется. Следовательно, в условиях опыта группа 1 сеянцев в среднем сильнее подвержена поражению коккомикозом по сравнению с группой 2.

Таблица А

Пары критических значений U_T (верхнее - подчеркнутое и ниже) для двустороннего критерия Манна-Уитни при уровне значимости $\alpha = 0,05$. (Рунион, 1982).

n_2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_1																				
1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
2	--	--	--	--	--	--	--	0	0	0	0	1	1	1	1	1	2	2	2	2
3	--	--	---	--	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	--	--	---	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	--	--	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	--	--	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	--	--	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	--	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	--	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	--	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	--	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	--	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	--	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	--	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	--	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	--	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	--	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	--	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	--	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	--	2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

(Прочерки в таблице указывают на невозможность принятия решения при установленном уровне значимости).

Критерий множественных сравнений Уилкоксона.

Этот критерий подобен критерию Манна-Уитни, но допускает число сравниваемых групп от 3 до 10. При этом предполагается, что комплекс равномерный, то есть количество наблюдений (n) во всех группах должно быть одинаково. Алгоритм вычисления этого критерия: 1) объединение всех выборок в одну, 2) их ранжирование в возрастающем порядке, 3) разделение по градациям фактора (условиям), 4) определение суммы рангов для каждого условия, 5) построение матрицы разностей сумм рангов для пар сравниваемых групп, 6) определение критической разности сумм и сравнение с ней парных разностей.

Пример. Необходимо сравнить четыре гибридные семьи алычи по экспертным оценкам (в баллах) хозяйственной ценности. В таблице представлена сумма оценок (баллов) по 11 признакам для четырех сравниваемых семей.

I	97	51	79	68	60	79	77	42	33	36	42
II	44	98	58	40	45	49	71	94	74	76	67
III	8	27	17	41	57	37	21	13	9	82	7
IV	46	90	75	81	93	81	83	99	70	65	53

Объединяем оценки всех четырех групп, располагаем их в возрастающем порядке и ранжируем (таблица).

От-метка	Ранг	Семья	От-метка	Ранг	Семья	От-метка	Ранг	Семья
7	1	III	45	16	II	76	31	II
8	2	III	46	17	IV	77	32	I
9	3	III	49	18	II	79	33,5	I
13	4	III	51	19	I	79	33,5	I
17	5	III	53	20	IV	81	35,5	IV
21	6	III	57	21	III	81	35,5	IV
27	7	III	58	22	II	82	37	III
33	8	I	60	23	I	83	38	IV
36	9	I	65	24	IV	90	39	IV
37	10	IV	67	25	II	93	40	IV
40	11	II	68	26	I	94	41	II
41	12	III	70	27	IV	97	42	I
42	13,5	I	71	28	II	98	43	II
42	13,5	I	74	29	II	99	44	IV
44	15	II	75	30	IV			

Перегруппируем данные в зависимости от принадлежности к исходным семьям и просуммируем ранги.

I	8	9	13,5	13,5	19	23	26	32	33,5	33,5	42	253
II	11	15	16	18	22	25	28	29	31	41	43	279
III	1	2	3	4	5	6	7	10	12	21	37	108
IV	17	20	24	27	30	35,5	35,5	38	39	40	44	350

Тест для проверки правильности расчетов:

$$R_1 + R_2 + R_3 + R_4 = kN/2(1 + kN)$$

R_1, R_2, R_3, R_4 – суммы рангов по группам (семьям)

k – число групп (семей)

N – количество наблюдений в каждой группе

$$253 + 279 + 108 + 350 = 990$$

$$4 \cdot 11/2(1 + 4 \cdot 11) = 22 \cdot 45 = 990$$

Формируем матрицу разностей сумм рангов:

		IV	II	I	III
		350	279	253	108
IV	350		71	97	242
II	279			26	171
I	253				145
III	108				

Для определения критического значения разности сумм обращаемся к специальной таблице Б. Табличное значение разности сумм рангов при числе признаков $n=11$, числе групп $k=4$ и уровне значимости $\alpha=0,05$ равно 155. Если какая-нибудь из наблюдаемых разностей превышает критическое значение или равна ему, то H_0 для этой пары семей отклоняется.

Вывод: с вероятностью 95% гибридная семья III является более ценной по комплексу признаков, чем семья IV и семья II. Все остальные различия недостоверны.

7.6. Метод максимального корреляционного пути

Этот метод позволяет выделить максимальные статистические связи между объектами матрицы. Он работает с матрицами таксономических отношений, коэффициентов корреляции т.п., и представляет собой так называемую неиерархическую кластер процедуру. В результате строится дендрит максимальных связей, который затем «разрезается» на кластеры или плеяды объектов (признаков).

Таблица Б

Критические разности в критерии Уилкоксона при сравнении пар градаций (групп генотипов, условий испытания и т. п.) для $k = 3, 4, \dots, 10$ и $n = 3, 4, \dots, 18, 20, 22, 24$. Наблюдаемая разность сумм значима при заданном уровне $\alpha = 0,05$ (светлый шрифт) или $\alpha = 0,01$ (жирный шрифт), если она равняется табличному значению или превышает его (Рунион, 1982).

n	k (число градаций)							
	3	4	5	6	7	8	9	10
3	15	23	30	37	45	52	60	68
	17	27	36	44	52	61	70	79
4	24	35	46	57	69	80	92	105
	27	42	54	67	80	94	107	121
5	33	48	63	79	96	112	129	146
	39	58	76	94	112	130	149	168
6	43	63	83	104	125	147	169	191
	51	76	99	123	147	171	196	221
7	54	79	105	131	158	185	213	241
	68	96	125	154	185	215	246	278
8	66	96	128	160	192	226	260	294
	82	117	152	188	225	263	301	339
9	79	115	152	190	229	269	310	351
	98	139	181	225	268	313	358	404
10	92	134	178	223	268	315	362	410
	115	163	212	263	314	366	420	473
11	106	155	205	257	309	363	418	473
	132	188	245	303	362	423	484	546
12	121	176	233	292	352	414	476	539
	150	214	278	345	413	481	551	621
13	136	199	263	329	397	466	537	608
	169	241	314	389	465	542	621	700
14	152	222	294	368	444	521	599	679
	189	269	351	434	519	606	694	783
15	169	246	326	408	492	577	665	753
	210	298	389	481	576	672	769	868
16	186	271	359	449	542	636	732	829
	231	328	428	530	634	740	847	956
17	203	296	393	492	593	696	802	908
	253	359	468	580	694	810	928	1047
18	221	323	428	536	646	759	873	989
	275	391	510	632	756	883	1011	1140
20	259	378	501	627	756	888	1022	1158
	322	458	597	740	886	1033	1183	1335
22	298	435	577	723	872	1024	1179	1336
	371	528	689	853	102	1192	1365	1540
24	340	496	657	824	994	1167	1343	1522
	422	601	784	972	1163	1358	1555	1754

Пример. Дана матрица парных коэффициентов ранговой корреляции между 7 признаками у абрикоса (1-окраска побега, 2-размер листьев, 3-толщина побега, 4-длина черешка, 5-окраска кожицы плода, 6-окраска мякоти плода, 7-окраска косточки). Необходимо построить так называемый максимальный корреляционный путь между признаками.

Матрица коэффициентов корреляции признаков у абрикоса

	1	2	3	4	5	6	7
1		-0,32	0,41	-0,19	0,74	0,02	0,13
2	-0,32		0,91	0,18	0,11	0,28	0,01
3	0,41	0,91		0,83	0,21	0,12	0,30
4	-0,19	0,18	0,83		-0,01	-0,03	0,40
5	0,74	0,11	0,21	-0,01		0,78	0,50
6	0,02	0,28	0,12	-0,03	0,78		0,80
7	0,13	0,01	0,30	0,40	0,50	0,80	

Сначала в данной матрице необходимо найти максимальное значение коэффициента (**0,91**). Поиск осуществляется, начиная с первой строки матрицы слева направо. Если максимальных значений в матрице несколько, дальнейший путь начинают с любого.

Поскольку в данной матрице представлены коэффициенты корреляции, следовательно, на всех этапах пути оперируют с абсолютными значениями этих коэффициентов. Далее строят вспомогательную таблицу.

Максимальный корреляционный путь (этап 1)

	1	2	3	4	5	6	7
2	${}_20,32^1$		${}_20,91^3$	${}_20,18^4$	${}_20,11^5$	${}_20,28^6$	${}_20,01^7$

Первой во вспомогательной таблице выписывается строка, содержащая максимальное значение коэффициента корреляции (то есть, строка 2), столбец, совпадающий с номером первой анализируемой строки в дальнейшем игнорируется, (то есть, столбец 2). В этой строке каждый коэффициент маркируется двумя индексами: номер строки (внизу слева) и номер столбца (вверху справа), например ${}_20,91^3$. Столбец, содержащий

максимальный коэффициент в дальнейшем игнорируется, (то есть, столбец 3).

Номер следующей выписываемой строки определяется номером столбца, содержащего максимальное значение коэффициента в предшествующей строке. В нашем примере следующая строка будет 3, так как максимальный коэффициент в матрице был ${}_2\mathbf{0,91}^3$.

При анализе очередной строки необходимо сравнивать коэффициент корреляции в соответствующей ячейке с коэффициентом в предыдущей строке этого же столбца и выбрать больший.

Максимальный корреляционный путь (этап 2)

	1	2	3	4	5	6	7
2	${}_2\mathbf{0,32}^1$		${}_2\mathbf{0,91}^3$	${}_2\mathbf{0,18}^4$	${}_2\mathbf{0,11}^5$	${}_2\mathbf{0,28}^6$	${}_2\mathbf{0,01}^7$
3	${}_3\mathbf{0,41}^1$			${}_3\mathbf{0,83}^4$	${}_3\mathbf{0,21}^5$	${}_3\mathbf{0,28}^6$	${}_3\mathbf{0,30}^7$

Анализируем строку 3. Значение коэффициента в 1-ом столбце равно ${}_3\mathbf{0,41}^1$, что больше, чем коэффициент в предыдущей строке 1-го столбца (${}_2\mathbf{0,32}^1$). Следовательно, выписываем это значение ${}_3\mathbf{0,41}^1$. Второй и третий столбцы игнорируются. Для 4 столбца значение коэффициента корреляции равно ${}_3\mathbf{0,83}^4$, а в предыдущей ${}_2\mathbf{0,18}^4$, следовательно, выбираем ${}_3\mathbf{0,83}^4$. Для 5 столбца значение коэффициента корреляции равно ${}_3\mathbf{0,21}^5$, а в предыдущей ${}_2\mathbf{0,11}^5$, следовательно, выписываем ${}_3\mathbf{0,21}^5$. Для 6 столбца значение коэффициента равно ${}_3\mathbf{0,28}^6$, что меньше, чем во 2-ой строке 6-го столбца (${}_2\mathbf{0,28}^6$). Значит, оставляем предыдущее значение ${}_2\mathbf{0,28}^6$. Для последнего 7 столбца значение коэффициента корреляции равно ${}_3\mathbf{0,30}^7$, что больше предыдущего ${}_2\mathbf{0,01}^7$, следовательно, выписываем ${}_3\mathbf{0,30}^7$. Переходим к сравнению коэффициентов 3-ей строки. Максимальным оказывается коэффициент ${}_3\mathbf{0,83}^4$, находящийся в 4 столбце, следовательно, следующей будет 4-ая строка (4 столбец в дальнейшем игнорируется).

Максимальный корреляционный путь (этап 3)

	1	2	3	4	5	6	7
2	${}_2\mathbf{0,32}^1$		${}_2\mathbf{0,91}^3$	${}_2\mathbf{0,18}^4$	${}_2\mathbf{0,11}^5$	${}_2\mathbf{0,28}^6$	${}_2\mathbf{0,01}^7$
3	${}_3\mathbf{0,41}^1$			${}_3\mathbf{0,83}^4$	${}_3\mathbf{0,21}^5$	${}_3\mathbf{0,28}^6$	${}_3\mathbf{0,30}^7$
4	${}_4\mathbf{0,19}^1$				${}_4\mathbf{0,21}^5$	${}_4\mathbf{0,28}^6$	${}_4\mathbf{0,40}^7$

Анализируем строку 4. В столбце 1 коэффициент равен ${}_4\mathbf{0,19}^1$, что меньше предыдущего ${}_3\mathbf{0,41}^1$, поэтому оставляем

значение предыдущего коэффициента ${}_30,41^1$. Второй, третий и четвертый столбцы игнорируем. В столбце 5 коэффициент равен ${}_40,01^5$, что меньше предыдущего ${}_30,21^5$, поэтому оставляем ${}_30,21^5$. В столбце 6, коэффициент равен ${}_40,03^6$, что меньше предыдущего ${}_20,28^6$, поэтому оставляем ${}_30,12^6$. В столбце 7 имеется коэффициент ${}_40,40^7$, который больше ${}_30,30^7$, поэтому выписываем ${}_40,40^7$. Таким образом, в строке 4 максимальным оказывается коэффициент ${}_30,41^1$, следовательно, следующей анализируемой строкой будет строка 1 (столбец 1 в последующем игнорируется).

Максимальный корреляционный путь (этап 4)

	1	2	3	4	5	6	7
2	${}_20,32^1$		${}_20,91^3$	${}_20,18^4$	${}_20,11^5$	${}_20,28^6$	${}_20,01^7$
3	${}_30,41^1$			${}_30,83^4$	${}_30,21^5$	${}_20,28^6$	${}_30,30^7$
4	${}_30,41^1$				${}_30,21^5$	${}_20,28^6$	${}_40,40^7$
1					${}_10,74^5$	${}_20,28^6$	${}_40,40^7$

Анализируем строку 1. Столбцы 1-4 игнорируем. В столбце 5 находится коэффициент ${}_10,74^5$, который больше предыдущего ${}_30,21^5$, следовательно, выписываем ${}_10,74^5$. В 6 и 7 столбцах коэффициенты равны соответственно ${}_10,02^6$ и ${}_10,13^7$, которые меньше предыдущих ${}_20,28^6$ и ${}_40,40^7$, поэтому эти два последних коэффициента остаются. Максимальным в строке 1 является коэффициент ${}_10,74^5$, поэтому следующей будет 5-ая строка (столбец 5 далее игнорируется).

Максимальный корреляционный путь (этап 5)

	1	2	3	4	5	6	7
2	${}_20,32^1$		${}_20,91^3$	${}_20,18^4$	${}_20,11^5$	${}_20,28^6$	${}_20,01^7$
3	${}_30,41^1$			${}_30,83^4$	${}_30,21^5$	${}_20,28^6$	${}_30,30^7$
4	${}_30,41^1$				${}_30,21^5$	${}_20,28^6$	${}_40,40^7$
1					${}_10,74^5$	${}_20,28^6$	${}_40,40^7$
5						${}_50,78^6$	${}_50,50^7$

Анализируем строку 5. Столбцы 1-5 игнорируем. В столбце 6 имеется коэффициент ${}_50,78^6$, который больше предыдущего ${}_20,28^6$, следовательно, он выписывается. Столбец 7 также содержит коэффициент ${}_50,50^7$ больший предыдущего ${}_40,40^7$. Следующей анализируется строка 6 (так как коэффициент ${}_50,78^6$ оказался максимальным), столбец 6 в дальнейшем игнорируется.

Максимальный корреляционный путь (этап 6)

	1	2	3	4	5	6	7
2	$_{2}0,32^1$		$_{2}0,91^3$	$_{2}0,18^4$	$_{2}0,11^5$	$_{2}0,28^6$	$_{2}0,01^7$
3	$_{3}0,41^1$			$_{3}0,83^4$	$_{3}0,21^5$	$_{3}0,28^6$	$_{3}0,30^7$
4	$_{3}0,41^1$				$_{3}0,21^5$	$_{4}0,28^6$	$_{4}0,40^7$
1					$_{1}0,74^5$	$_{2}0,28^6$	$_{4}0,40^7$
5						$_{5}0,78^6$	$_{5}0,50^7$
6							$_{6}0,80^7$

Анализируем строку 6. Столбцы 1-6 игнорируем. В столбце 7 имеется коэффициент $_{6}0,80^7$, который больше $_{5}0,50^7$, поэтому он выписывается. На этом путь завершен.

На основании полученных данных можно построить так называемый дендрит максимального корреляционного пути (рис. 29). Подстрочные и надстрочные индексы максимальных коэффициентов каждой строки (выделены жирным шрифтом) являются номерами объектов матрицы. Графическое их изображение показано на рис. 42.

После этого можно выделить плеяды сходных объектов. Разрезание максимального корреляционного пути для выделения плеяд проходит по наиболее слабой связи дендрита (связь между 3 и 1 признаками равная 0,41).

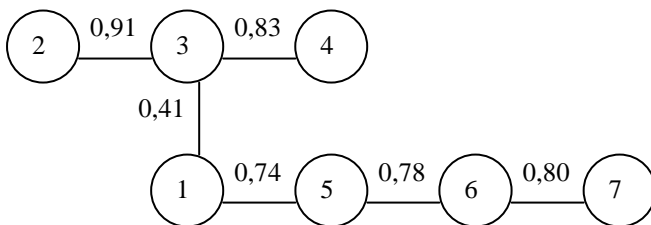


Рис. 42. Дендрит семи признаков абрикоса

В результате выделены две корреляционные плеяды признаков: плеяда 1 содержит три признака - 2,3,4; плеяда 2 содержит 4 признака - 1,5,6,7.

На следующем шаге рекомендуется определить средний коэффициент корреляции внутри каждой плеяды и сравнить его со средним коэффициентом корреляции между плеядами. Если внутрикластерный коэффициент корреляции достоверно превышает

межкластерный, то кластеризация проведена правильно. Если наоборот, то выбранный уровень разрезания максимальных связей дендрита был занижен и его следует увеличить.

Контрольные вопросы. 1. Каковы особенности многолетних культур как объектов исследования? 2. Какие типы шкал используются для описания признаков и в чем их особенности? 3. Чем отличаются одномерные математические модели и методы от многомерных? 4. Что называется рангом? 5. В каких случаях используют параметрические методы статистики, и в каких непараметрические? 6. Какие существуют типы статистических ошибок и как они связаны с понятием мощности критерия? 7. Какие имеются способы унификации признаков? 8. Какие непараметрические критерии используют при работе с номинальной и ранговой шкалами?

8. Словарь основных терминов и понятий

Аналитическая модель – **математическая модель** в виде уравнения (системы уравнений), если решение можно получить в явном виде.

Адекватная модель – модель, соответствующая данным, полученным в реальных экспериментах с объектом (системой, процессом).

Альтернативная гипотеза – утверждение, противоположное нулевой гипотезе.

Верификация модели – проверка **адекватности математической модели** на основе экспериментальных данных, не использованных для ее **идентификации и настройки**.

Дескриптивная модель – **аналитическая модель** для описания и прогнозирования поведения объекта.

Детерминистическая модель – модель, не учитывающая случайные неконтролируемые воздействия на моделируемый объект.

Динамическое программирование – математические методы оптимального планирования единой операции, состоящей из нескольких этапов, когда ситуации на очередных этапах зависят от управления объектом на предыдущих этапах операции.

Дискретная модель – модель, прогнозирующая процесс по «шкагам», в дискретные моменты времени.

Идентификация модели – подбор структуры уравнения (системы уравнений) для **аналитической модели**.

Имитационная модель - модель функционирования сложной системы, которую невозможно описать одной **аналитической моделью**; представляется сложным алгоритмом, реализуемым на компьютере.

Индекс – математическая формула, объединяющая в одну несколько **целевых функций многокритериальной задачи**.

Исследование операций – применение **математических моделей** и методов для обоснования решений во всех областях целенаправленной человеческой деятельности.

Коэффициент оригинальности – средняя сумма весов по присутствию и отсутствию **модальностей** каждого объекта исследуемой совокупности.

Линейное программирование – метод **оптимизации** для ситуаций, в которых **целевая функция** и ограничения линейны по **переменным управления**.

Математическая модель – система упрощенных предположений об объекте (системе, процессе), допускающих математическую

формализацию и применяемых, когда точные закономерности неизвестны или сложны.

Математическое программирование - группа математических методов, позволяющих выбрать из множества возможных **оптимальный план** воздействия на объект.

Медиана – состояние переменного, которое делит ранжированный ряд всех состояний на две одинаковые половины.

Методы оптимизации – математические методы, основанные на **оптимизационных моделях**.

Многокритериальная задача – оптимизационная задача с несколькими **целевыми функциями**.

Многомерная модель – модель для анализа объекта одновременно по нескольким переменным (признакам объекта).

Мода – **модальность**, встречающаяся в выборке с наибольшей частотой.

Модальность - состояние качественного признака в номинальной шкале.

Моделирование – это процесс построения и изучения модели объекта (системы, процесса).

Модель – это материальный или мыслительный объект, который по ходу изучения замещает объект – оригинал, сохраняя некоторые свойства последнего, важные для конкретного исследования.

Мощность статистического критерия – вероятность того, что критерий отвергнет ложную нулевую гипотезу.

Настройка модели – подбор числовых значений (констант), используемых в модели.

Нелинейное программирование – **метод оптимизации** для ситуаций, в которых либо **целевая функция**, либо ограничения, либо и то и другое нелинейны по **переменным управления**.

Непараметрические методы статистики – методы, которые не налагают модельных ограничений (предположений) на форму распределения анализируемых случайных величин.

Нулевая гипотеза – утверждение, которое исследователь обычно предполагает отклонить.

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

Оптимальное решение (план) – решение, по тем или другим критериям предпочтительное перед другими.

Оптимизационные модели – модели для подбора **оптимального плана** воздействий на объект.

Переменные управления – часть параметров объекта (системы, процесса) которыми можно управлять непосредственно.

Ранг - порядковый номер состояния переменного в ранжированном ряду всех его состояний.

Робастность – устойчивость математического метода к невыполнению модельных предположений.

Статистическая ошибка второго рода (β – типа) – случаи, когда принимают ложную нулевую гипотезу.

Статистическая ошибка первого рода (α – типа) – случаи, когда отклоняют верную нулевую гипотезу.

Стохастическая модель – модель, учитывающая случайные неконтролируемые воздействия на моделируемый объект.

Теория игр – метод моделирования конфликтных ситуаций для оптимизации поведения участников конфликта.

Унификация переменных – сведение оценок переменных к одной шкале.

Целевая функция – параметр моделируемого объекта, по которому оценивается эффективность плана воздействия на объект.

Шкала оценки – способ измерения состояния переменного.

9. Литература

Основная

1. Вентцель Е.С. Исследование операций: задачи, принципы, методология. М., Наука, 1988. 208 с.
2. Гильдерман Ю.И. Закон и случай. Новосибирск, Наука, 1991. 199 с.
3. Горстко А.Б. Познакомьтесь с математическим моделированием. М., Знание, 1991, 150 с.
4. Дворецкий Д.С., Дворецкий С.И., Муратова Е.И., Ермаков А.А. Компьютерное моделирование биотехнологических процессов и систем: Учеб. пособие. Тамбов: Изд. ТГТУ, 2005. 80 с.
5. Осипов Д.С. Математическое моделирование биосинтеза продуктов метаболизма М., 2002.
<http://www.studzona.com/referats/view/1542>
6. Смит Дж. Математические модели в биологии. М., Мир, 1970, 175 с.
7. Смит Дж. Модели в экологии. М., Мир, 1976, 184 с.
8. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере. М.: Инфра, 1997, 528 с.
9. Хеттманспергер Т. Статистические выводы, основанные на рангах. - М.: Финансы и статистика. 1987, 334 с.
10. Якушев В.П., Буре В.М. Статистический анализ опытных данных. Непараметрические критерии. Санкт-Петербург: АФИ, 2001, 61 с.

Дополнительная

1. Волова Т. Г., Войнов Н. А., Шишацкая Е. И., Калачева Г. С. Введение в биотехнологию. Версия 1.0 [Электронный ресурс]: методические указания по лабораторным работам. Красноярск: ИПК СФУ, 2008. 81 с.
2. Гильдерман Ю.И. Лекции по высшей математике для биологов. Новосибирск, Наука, 1974, 410 с.
3. Дегтярев Ю.П., Кузнецов Н.Г., Корниенко В.С., Коломок О.И. Математическое моделирование и оптимизация. Волгоград, Изд. ВГСХА, 1999, 218 с.

4. Матвеевко Е.Л., Мерзлов А.В., Довлетярова Э.А. Основы системного анализа и моделирования экосистем. М: Земля России, 2003, 720 с.
5. Моделирование микробной популяции. Лекция. <http://www.library.biophys.msu.ru/LectMB/lect11.htm>.
6. Протопопов И.И., Пашенко Ф.Ф. Компьютерное моделирование биотехнологических систем: Учеб. пособие. М.: МГУПБ, 2004. ч. 1. 123 с.
7. Ризниченко Г.Ю., Рубин А.Б. Математические модели биологии продукционных процессов. М., МГУ, 1993, 299 с.
8. Ризниченко Г.Ю. Лекции по математическому моделированию в биологии. Часть 1. М., МГУ., 2002, 230 с.
9. Рунион Р. Справочник по непараметрической статистике, М.: Финансы и статистика, 1982, 197 с.
10. Суховольский В.Г. Экономика живого. Новосибирск: Наука, 2004, 137 с.
11. Франс Дж., Торнли Дж. Математические модели в сельском хозяйстве. М., Агропромиздат, 1987, 400 с.

Содержание

	Введение.....	3
1.	Модели и моделирование.....	4
1.1.	Задачи моделирования.....	4
1.2.	Классификация моделей.....	5
1.3.	Значение моделирования.....	7
2.	Модели динамики биологических систем.....	9
2.1.	Прогрессия размножения.....	9
2.2.	Моделирование численности взаимодействующих популяций.....	11
2.3.	Модель баланса вещества и энергии.....	15
2.4.	Биологический метод борьбы с нежелательным видом.....	18
2.5.	Модель эпидемии.....	19
2.6.	Модели динамики возрастных групп.....	22
3.	Моделирование кинетики метаболизма в биотехнологии.....	26
3.1.	Кинетические характеристики процесса биосинтеза.....	27
3.2.	Моделирование кинетики биомассы.....	28
	Влияние концентрации питательной среды – субстрата.....	29
	Влияние концентрации основного продукта метаболизма.....	32
	Влияние отмирания микроорганизмов.....	34
3.3.	Моделирование кинетики образования продукта метаболизма.....	36
3.4.	Моделирование кинетики утилизации субстрата.....	39
3.5.	Моделирование накопления L-лейцина.....	39
4.	Вероятностные модели.....	43
4.1.	Сумма и произведение событий.....	44
4.2.	Формула полной вероятности.....	47
4.3.	Теория мишени.....	49
	Ряд Пуассона.....	51
	Приложения в экологии.....	54
	Редкие болезни, редкие признаки.....	56
5.	Исследование операций на основе оптимизационных моделей.....	59
5.1.	Линейное программирование.....	62
	Выбор курса лечения.....	63
	Рациональное размещение.....	66
	Определение плана перевозок (транспортная задача).....	68
5.2.	Нелинейное программирование.....	70
5.3.	Динамическое программирование.....	73
	Оптимизация пути.....	76
	Задача о распределении ресурсов.....	81
	Задача о загрузке машины.....	83
5.4.	Многокритериальные задачи.....	83
5.5.	Проблема оптимизации в условиях неопределенности.....	89

Теория игр.....	97
Игры с природой.....	103
5.6. Заключение: об исследовании операций вообще и в условиях неопределенности в частности.....	107
6. Имитационное моделирование.....	111
6.1. Принципы создания имитационной модели.....	111
6.2. Модели агробиоценоза.....	116
6.3. Модель сои.....	117
7. Применение непараметрических статистических моделей и методов на примере древесных культур.....	122
7.1. Особенности древесных культур как объектов моделирования.....	122
7.2. Шкалы измерений признаков.....	123
7.3. Унификация шкал признаков.....	128
7.4. Параметрические и непараметрические методы статистики.....	129
7.5. Алгоритмы и примеры вычисления непараметрических критериев.....	131
Номинальная шкала.....	131
Ранговая шкала.....	136
Критерий множественных сравнений Уилкоксона.....	139
7.6. Метод максимального корреляционного пути.....	140
8. Словарь основных терминов и понятий.....	147
9. Литература.....	150