

Міністерство аграрної політики та продовольства України  
Миколаївський національний аграрний університет

Кафедра економічної кібернетики і математичного моделювання

**Веселівська Н.В.**

***КУРС ЛЕКЦІЙ***

***з дисципліни: «Моделювання технологічних процесів і систем»  
спеціальність 8.09010101 «Агрономія»***

Миколаїв  
2014

## **Лекція 1. Основи математичного моделювання.**

### План

1. Поняття «модель» та «моделювання».
2. Етапи математичного моделювання.
3. Основні види та властивості моделей.
4. Принципи моделювання.
5. Пряма та зворотна задачі математичного моделювання.
6. Засоби математичного моделювання.
7. Основні методи розв'язання задач моделювання
8. Математична постановка задачі лінійного програмування.

*В мире не происходит ничего,  
в чем не был бы виден  
смысл какого-нибудь  
максимума или минимума.  
Л. Эйлер*

У сучасній науці істотно зростає роль математики і в розвитку сучасної агрономії. Майбутні спеціалісти потребують серйозної математичної підготовки, яка давала б можливість математичними методами вирішувати широке коло нових проблем, застосовувати обчислювальну техніку, використовувати теоретичні досягнення в практиці. Обробка експериментальних даних з використанням математичної статистики — це лише найбільш поширене, але не єдине і не найважливіше застосування математики. Врахувати взаємодію різноманітних факторів, що визначають структуру і особливості побутування природних систем, можна тільки за допомогою математичних методів і методів математичного моделювання. Важливим слід вважати процес побудови адекватної математичної моделі об'єкта або системи, що вивчається.

Математичне моделювання процесів, явищ та конструкцій складних технічних систем є могутнім інструментом аналізу величезної кількості задач. А імітація на ЕОМ функціонування складних (реально існуючих або проєктованих) систем — практично єдиний спосіб швидкого вирішення актуальних проблем науки та техніки.

Математизація великої кількості прикладних та фундаментальних наук дозволила моделювати поведінку частин системи, що розглядається, їх взаємодію з врахуванням факторів, що впливають на цю систему. Моделювання дозволяє значно спростити планування та виконання експериментів.

### **1. Поняття «модель» та «моделювання».**

Будь-які два об'єкти  $O_1$  та  $O_2$  людської діяльності в чомусь подібні, а в чомусь різні. Заміщення об'єкта  $O_1$  об'єктом  $O_2$  з метою вивчення найважливіших рис  $O_1$

за допомогою об'єкту O2 називається **моделюванням** об'єкта O1 об'єктом O2. При цьому об'єкт O1 називають **оригіналом** (або **натурою**), а об'єкт O2 – **моделлю**.

Таким чином, **модель** – це замісник оригіналу, який дозволяє вивчити або зафіксувати деякі важливі риси оригіналу.

Приклад 1. При проектуванні літаків, ракет, кораблів, автомобілів для дослідження залежності сили опору середовища рухові об'єкта від його швидкості (що практично неможливо визначити аналітично) використовують спосіб продувки зменшеної моделі транспортного засобу в аеродинамічній трубі. Заміщення оригіналу моделлю в цьому випадку дозволяє зафіксувати важливі властивості оригіналу та дослідити їх.

Приклад 2. При дослідженні роботи електричного генератора на споживача підключити реального споживача до генератора практично неможливо, тому споживача замінюють як правило послідовним підключенням опору та конденсатора. Тут має місце лише фіксація найбільш важливих рис оригіналу.

В загальному випадку процес моделювання складається з таких послідовних етапів:

- 1) Постановка задачі моделювання та визначення рис оригіналу, що підлягають дослідженню;
- 2) Вибір моделі, яка б фіксувала найбільш важливі риси оригіналу та яка б добре піддавалася дослідженню;
- 3) Дослідження моделі згідно поставленої задачі;
- 4) Перенесення результатів дослідження на оригінал;
- 5) Перевірка цих результатів.

Як моделі, що заміщують оригінали, можуть використовуватися зокрема засоби спілкування людей (мова, писемність), засоби пізнання матеріального світу (модель сонячної системи, модель будови атома), засоби навчання та тренування (тренажери для льотчиків, водіїв), засоби прогнозування поведінки оригіналів в різних умовах (наприклад, для підбору параметрів, що відповідають оптимальному функціонуванню роботи приладу в певних умовах).

В практичній діяльності інженера-електрика як в галузі виробництва, так і в галузі проектування або наукової діяльності в основному використовується моделювання з використанням методів та засобів математики, тому подальша увага буде сконцентрована саме на математичному моделюванні.

**Математичне моделювання** – це галузь науки і техніки, що забезпечує виявлення закономірностей проходження різноманітних явищ оточуючого світу або роботи систем та пристроїв шляхом їх математичного опису та моделювання без проведення натурних випробувань. При цьому використовують фундаментальні положення та закони математики, що описують явище, систему чи пристрій, що моделюються, на деякому рівні ідеалізації.

Таким чином, математична модель системи чи пристрою – це їх певний математичний опис, що забезпечує імітацію роботи системи чи пристрою на рівні, достатньо близькому до їх реальної поведінки при натурних випробуваннях.

## 2. Основні види та властивості моделей.

Моделі, що спрощують оригінал, зберігають подібність лише в самому істотному, називаються *гомоморфними*. Подібність в даному випадку однозначна лише в один бік: по моделі неможливо відтворити повністю оригінал, але вона дозволяє дослідити найбільш істотні його властивості. На межі гомоморфізм переходить в *ізоморфізм*. Ізоморфні моделі формально можна поміняти місцями, оскільки між ними існує взаємно однозначна відповідність.

Для розв'язання практичних задач недостатньо подібності. Необхідна можливість експериментувати на моделі. Це досягається перш за все спрощенням системи, обмеженням досліджуваних властивостей.

В залежності від способу відображення властивостей досліджуваної системи через ті чи інші носії всі моделі можна підрозділити на дві великі групи: *матеріальні (фізичні)* і *абстрактні*. По своїй природі фізичні моделі можуть бути механічними, електричними, гідравлічними і т.д. Фізичні моделі будуються на принципах прямої аналогії, коли оригінал і модель можуть відрізнитися лише масштабами, чи коли змінюються носії базових властивостей. Зустрічаються і змішані моделі. Вони припускають реальне втілення тих фізичних властивостей оригіналу, які цікавлять дослідника. Спрощені фізичні моделі, нерідко зменшених габаритів, називаються макетами. Тому фізичне моделювання часто іменують макетуванням. Абстрактні моделі, описують поведінку об'єктів абстрактно-логічними засобами, числовими, знаковими, графічними та інші.

**Математичні моделі.** Вони являють собою формалізовані опису об'єкта або системи за допомогою деякого абстрактного мови, наприклад у вигляді сукупності математичних співвідношень або схеми алгоритму. Розрізняють такі види математичного моделювання: вербальні (словесні), графічні, табличні, аналітичні та алгоритмічні. Нерідко математичні моделі виявляються придатними для опису безлічі систем та явищ в самих різних областях науки, техніки та економіки.

Класифікація математичних моделей багатоаспектна. Її можна здійснювати за різними ознаками і цілями, зокрема: за характером використання початкової інформації, за типом (видом) математичного методу, за мірою конкретизації моделюючого об'єкта тощо. Потрібно також розрізняти математичні моделі за особливостями опису ними просторових характеристик (властивостей) реальної системи.

В залежності від того, як враховуються випадкові фактори моделі поділяють на **детерміновані** (де усі параметри та впливи вважаються причиною обумовленими) та **стохастичні**, в яких важливу роль відіграють випадкові змінні (наприклад, в теорії фільтрації білий та гаусовський шуми). Стохастичні математичні моделі описують випадкові процеси, при цьому вважають, що випадковість тих чи інших явищ виражається в термінах ймовірності. З математичної точки зору детерміновані моделі є випадком стохастичних, з ймовірністю здійснення явища одиниця. З класу

стохастичних моделей використовують здебільшого регресійні моделі. Вони більш ефективні для вивчення не систем у цілому, а лише ситуацій у системі, де існує (або передбачається) однозначна відповідність між причиною і наслідком. Дуже часто в системах існує функціональний зв'язок між двома і більше змінними. Крім того, у більшості випадків функціональний зв'язок надзвичайно складний, або навіть зовсім невідомий. Тоді доводиться спиратися на певну гіпотезу щодо характеру функціональної залежності, тобто апроксимувати її деякими відносно простими математичними залежностями. Для пошуку таких залежностей між змінними використовують метод кореляційного та регресійного аналізу. Регресійний аналіз дає змогу побудувати, виходячи з масиву експериментальних даних, рівняння заданого виду. Кореляційний аналіз дає уявлення про те, наскільки тісно експериментальні точки узгоджуються з вибраним рівнянням.

За властивостями в перехідному режимі моделі прийнято поділяти на **статичні** та **динамічні**. В динамічних моделях розглядаються часово-змінні характеристики та взаємодії системи (на відміну від статичних систем, в яких ці параметри не залежать від часу).

В залежності від того, чи є в моделях керовані параметри, вони поділяються на **конструктивні** та **дескриптивні** (описуючі). На відміну від конструктивних описуючі моделі не містять керованих параметрів і тому не дозволяють вирішувати задачу про знаходження найбільш ефективного керування.

До найважливіших типів моделей можна також віднести **матричні** моделі. Назва цієї родини моделей означає, що кількість елементів структури і функцій системи обмежується певною таблицею чисел (матрицею). Моделі цього типу широко використовуються в екології для пошуку простих залежностей при відносно простих розрахунках (найчастіше популяційного характеру). Більш складні варіанти матричних моделей використовують для аналізу таких динамічних явищ, як кругообіг речовин та енергії на рівні балансу (вхід—вихід). Вадю цього типу моделей є прийняті положення про лінійність зв'язків між досліджуваними елементами природного середовища. Крім того, передбачається сталість структурних коефіцієнтів. В цілому при застосуванні цього типу моделей для екосистеми з'являється потреба у використанні високоагрегованих величин через нестачу багатьох фактичних даних або складність їх отримання. В кінцевому підсумку це може привести до грубих оцінок результатів моделювання.

На відміну від моделей, що характеризують випадкову зміну одного результативного параметра внаслідок дії кількох факторіальних, **багатовимірні** моделі враховують поведінку більш як одного результативного показника. Ці моделі особливо цінні для вивчення складних змінних системи, особливо складних взаємозв'язків між двома множинами. Багатовимірні моделі досить численні і придатні до розв'язування найрізноманітнішою класу задач (описових, оціночних, прогностичних) і не підходять до цілей оптимізації.

За ступенем математичної абстракції детерміновані моделі можна поділити на **складні** та **спрощені структури**. Складні математичні структури, що описують усі причинні зв'язки реальної системи, дозволяють доволі точно спрогнозувати поведінку системи в залежності від зміни параметрів. Спрощені структури, при яких вибирають ряд основних, найсуттєвіших залежностей, встановлюють і математично

описують зв'язки між окремими параметрами, що відповідають причинно-наслідковим закономірностям (ідеалізовані моделі).

Складна система може бути виражена комплексом з детермінованих, стохастичних та ін. моделей. Крім того, реальний процес протікає за обставин, що змінюються. Тому математичні моделі, що адекватно відображають дійсність в певний момент часу, можуть не відображати умови, що змінилися, в наступний момент. Постає потреба поновлення, адаптації моделей (тобто пристосування до параметрів, які змінилися), що виконується за допомогою ЕОМ.

Крім вищеназваних широких класів математичних моделей, існує ще багато інших класів.

Це, насамперед, **оптимізаційні** моделі. У загальному випадку оптимізаційні задачі можуть бути поділені на три групи: детерміновані, стохастичні та задачі, що вирішуються в умовах невизначеності (або неповної інформації). Останні характеризуються тим, що стохастичні характеристики системи, яка досліджується, не можуть бути отримані. Існуючі методи розв'язування цих задач, як правило, враховують тільки доволі мані зміни коефіцієнтів цільової функції та системи обмежень моделі, і практично не дозволяють врахувати варіації структури моделі. Оптимізаційна задача вимагає переробки великої кількості експериментального матеріалу, його треба проаналізувати та на основі отриманих результатів прийняти рішення для подальшого удосконалення процесу.

Останнім часом успішно намагаються побудувати екологічні моделі, які спираються на теорію катастроф. Моделі цього типу застосовують щодо систем, які мають властивості бімодальності, перервності, гістерезису і дивергенції, а також характеризуються диссипативністю енергетичних та матеріальних потоків. Такими властивостями характеризуються, наприклад, екосистеми ґрунту, їх біогеохімічні цикли та енергетичні потоки. Для екосистеми ґрунту характерним є перебування її у багатьох станах, що пов'язано з добовими, сезонними, річними і багаторічними циклами природної динаміки (бімодальність). Екосистеми ґрунту уповільнено реагують на сторонні впливи (за рахунок потужного механізму гомеостазу)(гістерезис). Ґрунтові системи функціонують так, що близькі вихідні умови нерідко еволюціонують у досить віддалені кінцеві результати (дивергенція). Складніше описати перервність екосистеми ґрунту і процесів, що відбуваються, оскільки ґрунт, як природне тіло, належить до континуальних об'єктів природи.

Широко застосовуються моделі систем, що розвиваються. Їх основною властивістю є така: на момент початку розвитку присутні деякі початкові ресурси. У динамічну систему повинна надходити речовина, енергія та інформація. У системах розвитку повинна бути підсистема відтворення та вдосконалення, має враховуватися характер використання умов довкілля, у взаємодії з яким система створює та поглинає продукти, а також виділяє застарілі та непотрібні. Повинні виконуватися деякі балансові співвідношення між субстратами, що надходять у динамічну систему, та продуктами динамічної системи (причому обов'язковим є функціональний взаємозв'язок між ресурсами, що витрачаються на внутрішній розвиток та на виконання зовнішніх функцій динамічної системи, між швидкістю відтворення ресурсів, інтенсивністю їх використання та результатами функціонування системи).

При програмуванні урожайності сільськогосподарських культур важливе значення має розробка методу багатокритеріальної оптимізації економічних функцій розподілу ресурсів, що базуються на адекватних залежностях „урожайність – комплекс факторів” [1,3], за якими розраховують урожайність при певній величині добрив, води, рівня агротехніки, тощо [2,4].

При програмуванні урожайності необхідно враховувати адекватну виробничу функцію залежності врожаю від різних факторів. При цьому необхідно оптимізувати і іншу функцію - функцію вартості ресурсів, які при цьому витрачаються. Тому в даній роботі ставиться задача двохкритеріальної оптимізації рівнів „вода-добрива” з такого розрахунку, щоб одержувати урожайність при мінімізації вартості ресурсів [3].

Розглянемо самоорганізаційні моделі. На основі принципу самоорганізації моделей на ЕОМ здійснюється перебір моделей різної складності за рядом критеріїв та знаходження оптимальної. За обраними критеріями визначається модель оптимальної структури.

В математичних моделях окремо виділяють **імітаційні** моделі. Термін «імітаційне моделювання» означає, що маємо справу з такими математичними моделями, за допомогою яких результат неможливо попередньо обчислити або передбачити (для передбачення поведінки реальної системи необхідно проводити її експеримент, імітацію, на моделі за заданих початкових даних). Імітація являє собою чисельний метод проведення на ЕОМ експериментів з математичними моделями, які описують поведінку системи впродовж заданого або сформованого періоду часу. Поведінка компонент складної системи та їх взаємодія в імітаційному моделюванні часто описуються набором алгоритмів (реалізуються певною мовою моделювання). Усі ці описи є програмною імітацією, яку необхідно спочатку підладити та випробувати, а потім використати для постановки імітаційного експерименту на ЕОМ. Тому під процесом імітації на ЕОМ розуміють конструювання моделі, її випробування та використання для вивчення певного явища або проблеми. Отже, імітаційні моделі є достатньо складним програмним виробом. Технологія їх проектування — працемісткий процес, що включає методи та засоби, які забезпечують створення та розвиток моделі. Великі розміри системи, складність поведінки і складників, велика вартість розробки вимагають математичного моделювання на всіх етапах проектування такої системи. Тому моделювання супроводжує і процес проектування, і процес випробування, і процес експлуатації складної системи.

### **Основні властивості моделей.**

Моделі мають ряд властивостей, від яких залежить успіх їх застосування в практиці моделювання. Відзначимо лише деякі з них, найбільш важливі:

Адекватність - це ступінь відповідності моделі досліджуваного реальному об'єкту. Вона ніколи не може бути повною. На практиці модель вважають

адекватною, якщо вона з задовільною точністю дозволяє досягти цілей дослідження.

**Простота (складність)** - чим більша кількість властивостей об'єкта описує модель, тим більш складною вона виявляється. Не завжди чим складніше модель, тим вище її адекватність. Треба прагнути знайти найбільш просту модель, що дозволяє досягти бажаних результатів вивчення.

**Потенційність** (самий корінь) - здатність моделі дати нові знання про досліджуваний об'єкт, спрогнозувати його поведінку або властивості. На основі вивчення математичних моделей, що описують рух планет Сонячної системи з урахуванням закону всесвітнього тяжіння, теоретично були передбачені існування орбіти планет Нептун і Плутон.

#### 4. Принципи моделювання.

Моделювання базується на декількох основоположних принципах.

**Принцип інформаційної достатності.** При повній відсутності інформації про досліджуваний об'єкт побудова його моделі неможливо. З іншого боку, за наявності повної інформації про об'єкт побудова його моделі не має сенсу. Існує певний рівень апріорної інформації про об'єкт, при досягненні якої може бути побудована його адекватна модель.

**Принцип здійсненності.** Створювана модель повинна забезпечувати досягнення поставленої мети досліджуваного з імовірністю, істотно відрізняється від нуля.

**Принцип множинності моделей.** Даний принцип є ключовим. Мова йде про те, що створювана модель повинна відображати в першу чергу ті властивості реальної системи, які цікавлять дослідника. Відповідно, при використанні будь-якої конкретної моделі пізнаються лише деякі сторони реальності. Для більш повного її дослідження необхідний ряд моделей, що дозволяє з різних сторін і з різним ступенем деталізації розглянути досліджуваний об'єкт.

**Принцип агрегування.** У більшості випадків складну систему можна уявити що складається з агрегатів (підсистем), для адекватного математичного опису яких виявляються придатним деякі стандартні математичні схеми.

**Принцип параметризації.** Цей принцип означає, що модель будується у вигляді відомої системи, параметри якої невідомі.

#### 5. Пряма та зворотна задачі математичного моделювання.

**Пряма задача:** всі параметри системи, що досліджується, відомі, досліджується поведінка моделі в різних умовах.

**Зворотні задачі:**

- а) **Задача розпізнавання:** визначення параметрів моделі шляхом співставлення даних спостереження та результатів моделювання. За результатами спостережень намагаються визначити, які процеси



управляють поведінкою об'єкта, знаходять визначальні параметри моделі. В зворотній задачі розпізнавання треба визначити значення параметрів моделі за відомою поведінкою системи як цілого.

Приклади задач розпізнавання:

- Задача електророзвідки: визначення підземних структур за допомогою вимірювань на поверхні.
  - Задача магнітної дефектоскопії: визначення дефекту в деталі, розташованій між полюсами магніту, за збуреннями магнітного поля на поверхні деталі.
- б) Задача синтезу (задача математичного проектування): побудова математичних моделей систем та пристроїв, які повинні мати задані технічні характеристики. На відміну від задач розпізнавання, що полягають у визначенні параметрів моделі, відповідній реальному стану системи, в задачах синтезу відсутня вимога унікальності рішення. Це дозволяє з кількох можливих рішень обрати найбільш прийнятний результат.

## **6. Засоби математичного моделювання.**

Математичне моделювання суспільних, економічних, біологічних та фізичних явищ, об'єктів, систем та різних пристроїв – важливий засіб пізнання природи та проектування доволі різноманітних систем та пристроїв. Хрестоматійними стали приклади ефективного використання моделювання в створенні ядерних технологій, авіаційних та аерокосмічних систем, в прогнозуванні атмосферних та океанічних явищ, погоди тощо.

Однак для таких сфер моделювання частіше за все потрібні суперкомп'ютери та роки роботи великих груп вчених для підготовки даних для моделювання та його відпрацювання.

Між тим математичне моделювання на рівні рішення більш простих задач, наприклад, в галузі механіки, електротехніки, автоматики, електроніки та радіотехніки (та ще багатьох галузей науки та техніки) як потреб виробництва, так і для потреб проектування чи наукових досліджень в наш час стало доступним багатьом користувачам сучасних ПК. А при використанні узагальнених моделей стає можливим моделювання досить складних систем, наприклад систем управління, електроенергетичних систем або промислових комплексів.

Існує досить багато програмних пакетів, які дозволяють більш легко проводити моделювання різноманітних систем та пристроїв. Для здійснення математичних досліджень та розрахунків зокрема використовуються пакети Mathematica, Maple, Mathcad, для моделювання електронних пристроїв – EWB, для моделювання систем управління та електротехнічних систем – VisSim, для наукових досліджень та для програмування роботи технічних систем – LabView, але найбільше розповсюдження практично у всіх прикладних галузях отримав програмний пакет MATLAB, що має в своєму складі багато спеціалізованих за окремими напрямками чи галузями програмних пакетів. Як складова пакету MATLAB для проведення моделювання різноманітних систем використовується Simulink – система наглядного моделювання, яка і буде використовуватися при проведенні лабораторних робіт.

## 7. Основні методи розв'язання задач моделювання

На етапі програмної реалізації моделі та реалізації плану експериментів необхідний вибір методів рішення задач моделювання. При цьому використовуються три основні групи методів:

- 1) Графічні - оціночні наближені методи, засновані на побудові та аналізі графіків;
- 2) Аналітичні - рішення, отримані суворо у вигляді аналітичних виразів (Придатні для вузького кола завдань);
- 3) Чисельні - основний інструмент для вирішення складних математичних задач, заснованих на застосуванні різних чисельних методів.

## ЛІТЕРАТУРА

1. Математическое моделирование экономических процессов в сельском хозяйстве / Гатаулин А.М., Гаврилов Г.В., Сорокина Т.Н. и др.; Под ред. А.М. Гатаулина. - М.: Агропромиздат, 1990, - 432 с.
2. Доспехов А.М. Методика полевого опыта (с основами статистической обработки результатов исследований). – 5-е изд., доп. и перераб. – М.: Агропромиздат, 1985 - 351 с.
3. Ковальчук П.І., Лазер П.Н., Пендак Н.В., Матяш Т.В., Зябров І.А.. Інтегроване управління водними, земельними та матеріальними ресурсами у зрошуваному землеробстві.// Таврійський науковий вісник: Збірник наукових праць ХДАУ. Вип.52. – Херсон: Айлант.2007.-с.87-92.
4. Вергунова І.М. Основи математичного моделювання для аналізу та прогнозу агрономічних процесів. – К.:Нора-Прінт, 2000. – 146 с.
5. Вергунов В.А., Вергунова И.Н., Шкрабак В.С. Основы математического моделирования для анализа и прогноза агрономических процессов.- Типография С-ПбГАУ/ООО «Литера», 2003.- 219 с.
6. Компьютерное моделирование биотехнологических процессов и систем: Учеб. пособие / Д.С. Дворецкий, С.И. Дворецкий, Е.И. Муратова, А.А. Ермаков. Тамбов: Изд-во Тамб. гос. техн. ун-та, 2005. 80 с.
7. Леснікова І.Ю., Харченко Є.М. Основи роботи і вирішення задач сільського господарства в середовищі електронних таблиць EXCEL. 2002.- 145 с.

## Лекція №2

### Графічний метод розв'язування задач лінійного програмування

Для розв'язування двовимірних задач лінійного програмування, тобто задач із двома змінними, а також деяких тривимірних задач застосовують графічний метод, що ґрунтується на геометричній інтерпретації та аналітичних властивостях задач лінійного програмування. Обмежене використання графічного методу зумовлене складністю побудови багатогранника розв'язків у тривимірному просторі (для задач з трьома змінними), а графічне зображення задач з кількістю змінних більше трьох взагалі неможливе.

Розглянемо задачу.

Знайти

$$\max(\min) Z = c_1 x_1 + c_2 x_2 \quad (2.17)$$

за умов:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 (\leq, =, \geq) b_1; \\ a_{21}x_1 + a_{22}x_2 (\leq, =, \geq) b_2; \\ a_{31}x_1 + a_{32}x_2 (\leq, =, \geq) b_3; \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 (\leq, =, \geq) b_m. \end{cases} \quad (2.18)$$

$$x_1 \geq 0, x_2 \geq 0 \quad (2.19)$$

Припустимо, що система (2.18) за умов (2.19) сумісна і багатокутник її розв'язків обмежений.

Згідно з геометричною інтерпретацією задачі лінійного програмування (§ 2.4) кожне  $i$ -те обмеження-нерівність у (2.18) визначає півплощину з граничною прямою  $a_{i1}x_1 + a_{i2}x_2 = b_i$  ( $i = 1, 2, \dots, m$ ). Системою обмежень (2.18) графічно можна зобразити спільну частину, або переріз усіх зазначених півплощин, тобто множину точок, координати яких задовольняють всі обмеження задачі — *багатокутник розв'язків*.

Умова (2.19) невід'ємності змінних означає, що область допустимих розв'язків задачі належить першому квадранту системи координат двовимірного простору. Цільова функція задачі лінійного програмування геометрично інтерпретується як сім'я паралельних прямих  $c_1x_1 + c_2x_2 = \text{const}$ .

Скористаємося для графічного розв'язання задачі лінійного програмування властивостями, наведеними в § 2.5:

якщо задача лінійного програмування має оптимальний план, то екстремального значення цільова функція набуває в одній із вершин її багатокутника розв'язків. Якщо ж цільова функція досягає екстремального значення більш як в одній вершині багатокутника, то вона досягає його і в будь-якій точці, що є лінійною комбінацією цих вершин.

Отже, розв'язати задачу лінійного програмування графічно означає знайти таку вершину багатокутника розв'язків, у результаті підстановки координат якої в (2.17) лінійна цільова функція набуває найбільшого (найменшого) значення.

**Алгоритм графічного методу** розв'язування задачі лінійного програмування складається з таких кроків:

1. Будуємо прямі, рівняння яких дістаємо заміною в обмеженнях задачі (2.18) знаків нерівностей на знаки рівностей.
2. Визначаємо півплощини, що відповідають кожному обмеженню задачі.
3. Знаходимо багатокутник розв'язків задачі лінійного програмування.
4. Будуємо вектор  $\vec{N} = (c_1; c_2)$ , що задає напрям зростання значення цільової функції задачі.

5. Будуємо пряму  $c_1x_1 + c_2x_2 = \text{const}$ , перпендикулярну до вектора  $\vec{N}$ .

6. Рухаючи пряму  $c_1x_1 + c_2x_2 = \text{const}$  в напрямку вектора  $\vec{N}$  (для задачі максимізації) або в протилежному напрямку (для задачі мінімізації), знаходимо вершину багатокутника розв'язків, де цільова функція набуває екстремального значення.

7. Визначаємо координати точки, в якій цільова функція набуває максимального (мінімального) значення, і обчислюємо екстремальне значення цільової функції в цій точці.

У разі застосування графічного методу для розв'язування задач лінійного програмування можливі такі випадки:

1. Цільова функція набуває максимального значення в єдиній вершині  $A$  багатокутника розв'язків (рис. 2.5).

2. Максимального значення цільова функція досягає в будь-якій точці відрізка  $AB$  (рис. 2.6). Тоді задача лінійного програмування має альтернативні оптимальні плани.

3. Задача лінійного програмування не має оптимальних планів: якщо цільова функція необмежена згори (рис. 2.7) або система обмежень задачі несумісна (рис. 2.8).

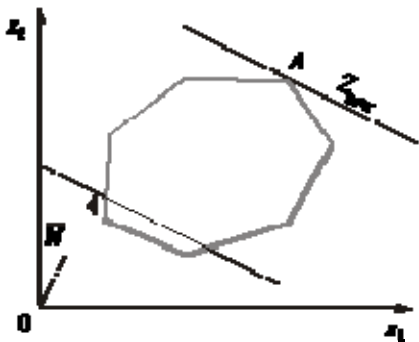


Рис. 2.5

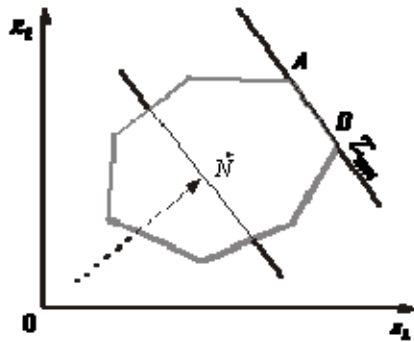


Рис. 2.6

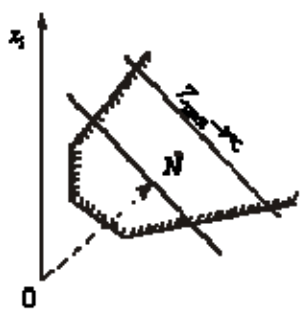


Рис. 2.7

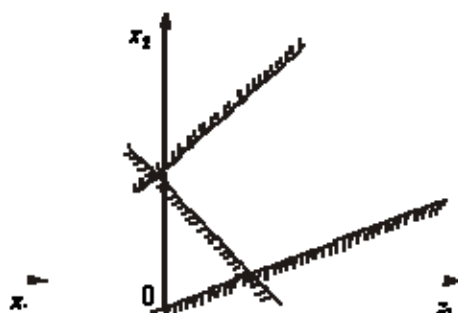


Рис. 2.8

4. Задача лінійного програмування має оптимальний план за необмеженої області допустимих розв'язків (рис. 2.9 і 2.10). На рис. 2.9 у точці  $B$  маємо максимум, на рис. 2.10 у точці  $A$  — мінімум, на рис. 2.11 зображено, як у разі необмеженої області допустимих планів цільова функція може набирати максимального чи мінімального значення у будь-якій точці променя.

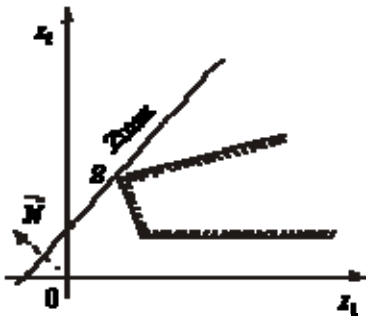


Рис. 2.9

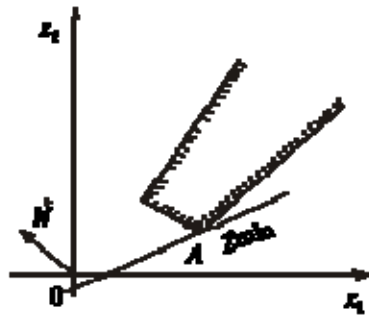


Рис. 2.10

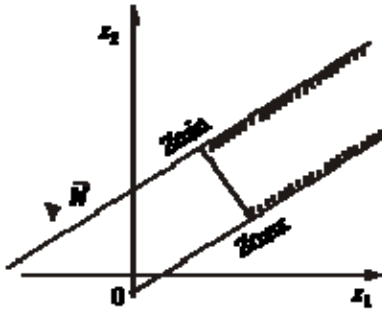


Рис. 2.11

Розв'язувати графічним методом можна також задачі лінійного програмування  $n$ -вимірному просторі, де  $n > 3$ , якщо при зведенні системи нерівностей задачі до системи рівнянь шляхом введення додаткових змінних кількість змінних  $n$  на дві більша, ніж число обмежень  $m$ , тобто  $n - m = 2$ .

Тоді, як відомо з курсу вищої математики, можна дві з  $n$  змінних, наприклад  $x_1$  та  $x_2$ , вибрати як вільні, а інші  $m$  зробити базисними і виразити через вільні. Припустимо, що це зроблено. Отримаємо  $m = n - 2$  рівнянь вигляду:

$$\begin{cases} x_3 = \alpha_{31}x_1 + \alpha_{32}x_2 + \beta_3; \\ x_4 = \alpha_{41}x_1 + \alpha_{42}x_2 + \beta_4; \\ \dots \\ x_m = \alpha_{m1}x_1 + \alpha_{m2}x_2 + \beta_m. \end{cases}$$

Оскільки всі значення  $x_i \geq 0$  ( $i = \overline{1, n}$ ), то мають виконуватись умови:

$$x_1 \geq 0, x_2 \geq 0,$$

$$\begin{cases} x_3 = \alpha_{31}x_1 + \alpha_{32}x_2 + \beta_3 \geq 0; \\ x_4 = \alpha_{41}x_1 + \alpha_{42}x_2 + \beta_4 \geq 0; \\ \dots \\ x_m = \alpha_{m1}x_1 + \alpha_{m2}x_2 + \beta_m \geq 0. \end{cases} \quad (2.19.1)$$

Розглянемо, як можна зобразити ці умови геометрично. Візьмемо, наприклад, першу з них:

$$x_3 - \alpha_{31}x_1 + \alpha_{32}x_2 + \beta_3 \geq 0.$$

Узявши величину  $x_3$  рівною своєму крайньому значенню — нулю, отримаємо рівняння:

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

Це рівняння прямої. Для такої прямої  $x_1 = 0$ , по одну сторону від неї  $x_1 > 0$ , а по другу —  $x_1 < 0$ . Відмітимо ту сторону прямої  $a_{31}x_1 + a_{32}x_2 + \beta_3 = 0$ , де  $x_3 > 0$ .

В аналогічний спосіб побудуємо і всі інші обмежуючі прямі:  $x_4 = 1$ ;  $x_1 = 0$ ; ...;  $x_n = 0$  і відмітимо для кожної з них півплощину, де відповідна змінна більше нуля.

У такий спосіб отримують  $n - 2$  прямі та дві осі координат ( $x_1 = 0$ ,  $x_2 = 0$ ). Кожна з них визначає півплощину, де виконується умова  $x_i > 0 (i = \overline{1, n-2})$ . Частина площини в  $x_1, x_2$  належить водночас всім півплощинам, утворюючи багатокутник допустимих розв'язків.

Припустимо, що в задачі необхідно знайти максимальне значення функціонала:

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

Підставивши вирази для  $x_3, x_4, x_5, \dots, x_n$  з (2.19.1) у цей функціонал, зведемо подібні доданки і отримаємо вираз лінійної функції  $F$  всіх  $n$  змінних лише через дві вільні змінні  $x_1$  та  $x_2$ :

$$F = \gamma_0 + \gamma_1x_1 + \gamma_2x_2,$$

де  $\gamma_0$  — вільний член, якого в початковому вигляді функціонала не було.

Очевидно, що лінійна функція  $F = \gamma_1x_1 + \gamma_2x_2$  досягає свого максимального значення за тих самих значень  $x_1$  та  $x_2$ , що й  $F = \gamma_0 + \gamma_1x_1 + \gamma_2x_2$ . Отже, процедура відшукування оптимального плану з множини допустимих далі здійснюється за алгоритмом для випадку двох змінних.

### Лекція 3.

#### Економіко-математична модель оптимізації структури посівних площ.

##### *Загальна постановка задачі*

В сільському господарстві земля - основний засіб виробництва. Її необхідно використовувати з максимальним економічним ефектом, постійно дбаючи про родючість. Успішне рішення задачі раціонального використання землі багато в чому залежить від обґрунтованої структури посівних площ.

Оптимальна структура посівних площ являє собою відображення спеціалізації рослинництва у вигляді такого співвідношення посівних площ по культурах, яке забезпечує отримання максимальної кількості сільськогосподарської продукції.

Постановка задачі: виходячи з виробничих ресурсів (земельних, трудових, матеріальних, тощо) визначити оптимальну структуру посівних площ, яка забезпечить виконання плану продажу продукції по видах, внутрішні потреби господарства в продукції, агрономічні та сівозмінні вимоги при максимальному економічному ефекті.

В якості критерію оптимальності може бути:

- максимум отримання прибутку від виробництва;
- максимум виробництва валової та товарної продукції в грошовому виразі;
- максимум виробництва конкретної продукції (зерна, картоплі, тощо).

Для розробки економіко-математичної моделі необхідна така інформація:

- розмір площі ріллі, сінокосів, пасовищ;
- перелік сільськогосподарських культур, що вирощуються в даній кліматичній зоні, їх можлива врожайність;
- затрати праці та коштів на 1га посіву;
- виручка від реалізації продукції (з 1га або одиниці продукції);
- наявність виробничих ресурсів в господарстві, норми витрат їх на 1га;
- потреба тварин в кормах по видах;
- план продажу та внутрішня потреба господарства по видах продукції;
- агротехнічні вимоги та можливі межі входження окремої культури або групи культур в сівозміну.

Економіко-математичну модель можна представити в загальному виді.

Знайти максимум функції:

$$Z_{\max} = \sum_{j=1}^n C_j X_j$$

при обмеженнях:

- виробничі ресурси господарства обмежені:

$$\sum_{j=1}^n a_{ij} X_j \leq b_i;$$

- окремі культури та групи культур в сівозміну вводяться в обґрунтованих агротехнічними вимогами межах:

$$\overline{X1}_j \leq \sum_{j=1}^n X_j \leq X2_j ;$$

- дотримується співвідношення окремих культур (наприклад: для отримання збалансованих концентрованих кормів, страхові культури, ...):

$$\sum_{j=1}^l W_{ij} X_j - \sum_{j=l+1}^n W'_{ij} X_j \leq 0;$$

- обмеження, на отримання продукції даного виду не менше заданої кількості:

$$\sum_{j=1}^n U_{ij} X_j \geq P_i;$$

- при умові невід'ємності змінних:

$$X_j \geq 0; \overline{X}_j \geq 0;$$

де:

$C_j$  - прибуток з 1га  $j$ -ї культури;

$X_j$  - площа посіву  $j$ -тої культури;

$a_{ij}$  - затрати ресурсів  $i$ -го виду на 1га  $j$ -ї культури;

$b_i$  - наявність виробничих ресурсів  $i$ -го виду;

$\overline{X1}_j$  - нижня межа входження окремої культури або групи культур в сівозміну;

$\overline{X2}_j$  - верхня межа входження окремої культури або групи культур в сівозміну;

$U_{ij}$  - урожайність  $j$ -тої культури;

$P_i$  - необхідна кількість  $i$ -го виду продукції;

$W_{ij}, W'_{ij}$  - коефіцієнти пропорційності.

### Умова задачі

Господарство має 1200га оранки, на яких може вирощувати зернові та просапні культури.

Площа зернових у структурі посівних площ повинна бути від 40 до 60%, площа цукрових буряків не повинна перевищувати 15%, а площа соняшника або картоплі – не більше 10%. Площа озимих повинна бути не більше 50% від максимальної площі зернових. Площа ярої пшениці співвідноситься до площі озимої пшениці як 1:10.

Підприємство повинно виробити не менше:

- 15600ц зерна (в тому числі 7050ц пшениці);
- 10500ц цукрового буряку;
- 1050ц соняшника;
- 530ц проса або гречки.

Крім цього для галузі тваринництва необхідно:

- соломи – 17400ц;
- кукурудзи на силос (зелений корм) – 15570ц

В якості критерію оптимальності виступає прибуток від товарної продукції рослинництва.

Культури для розробки економіко-математичної моделі наведено в таблиці 7, вихідні дані взяти довільно з заданого діапазону (таблиця 1).

Розробити економіко-математичну модель, змінними в якій будуть площі під кожен культуру ( $X_1, X_2, \dots, X_8$ ), валові збори культур ( $X_9, X_{10}, \dots, X_{16}$ ), загальні витрати на виробництво продукції ( $X_{17}$ ), витрати на виробництво товарної продукції ( $X_{18}$ ), виручка від реалізації товарної продукції рослинництва ( $X_{19}$ ).

Таблиця 1

Культури, що будуть висіватися в господарстві

Варіант	Зернові					Просапні		
	озима пшениця	яра пшениця	ячмінь	гречка	просо	цукровий буряк	соняшник	кукурудза на з/к
<b>3</b>								

Таблиця 2

Вихідні дані для розробки економіко-математичної моделі

Культура	Урожайність, ц/га		Затрати на 1 га, грн.	Ціна реалізації 1ц, грн.
	основна продукція	побічна продукція		
Озима пшениця	45	55	600	40
Яра пшениця	30	30	480	40
Гречка	12		480	44
Ячмінь	35	35	300	38
Просо	20	25	450	29
Цукровий буряк	280		2300	16
Соняшник	23		650	90
Кукурудза на з/к	250		300	

### Рішення

На 19-ть змінних задачі накладено 25 обмежень:



1-ше обмеження стосується використання наявної площі:

$$1X_1+1X_2+1X_3+1X_4+1X_5+1X_6+1X_7+1X_8 \leq 1200, \text{ де:}$$

$X_1, X_2, \dots, X_8$  – площі сільськогосподарських культур.

2-ге та 3-тє обмеження задають верхню та нижню межу насиченості сівозміни зерновими культурами, питома вага яких повинна бути від 40% до 60%.

$$1X_1+1X_2+1X_3+1X_4+1X_5 \leq 720 \quad (60\%)$$

$$1X_1+1X_2+1X_3+1X_4+1X_5 \geq 480 \quad (40\%), \text{ де}$$

$X_1, X_2, X_3, X_4, X_5$  – площі зернових культур.

4-те, 5-те та 6-те обмеження описують максимальну площу цукрових буряків, озимих та картоплі відповідно, так 4-те стосується площі цукрових буряків:

$$1X_6 \leq 180 \quad (15\%), \text{ де:}$$

$X_6$  – площа цукрових буряків.

7-ме обмеження характеризує співвідношення між площею посіву озимої пшениці та ярої пшениці:

$$1X_1 - 10X_2 = 0, \text{ де:}$$

$X_1$  – площа озимої пшениці;

$X_2$  – площа ярої пшениці.

Група обмежень з 8-го по 12-те задають нижню межу виробництва певної продукції, так 8-ме обмеження характеризується валового збору зерна, який повинен бути на менше 15600 ц:

$$1X_9+1X_{10}+1X_{11}+1X_{12}+1X_{13} \geq 15600; \text{ де:}$$

$X_9, X_{10}, X_{11}, X_{12}, X_{13}$  – валові збори озимої пшениці, ярої пшениці, гречки, жита, гороху відповідно.

13-те та 14-те обмеження забезпечать задоволення потреб тваринництва в соломі та силосі відповідно, так 13-те має вигляд:

$$55X_1+30X_2+35X_4+25X_5 \geq 17400; \text{ де:}$$

$X_1, X_2, X_4, X_5$  – площі зернових культур: озимої пшениці, ярої пшениці, жита та гороху відповідно (коефіцієнти при цих змінних - урожайність соломи відповідних культур).

Обмеження з 15-го по 22-ге забезпечують зв'язок між площею посіву та валовим збором по кожній сільськогосподарській культурі, 15-те обмеження стосується озимої пшениці:

$$45X_1 - X_9 = 0; \text{ де:}$$

$X_1$  – площа озимої пшениці (32 – урожайність озимої пшениці);

$X_9$  – валовий збір озимої пшениці.

23-те обмеження акумулює затрати на виробництво продукції:

$$600X_1+480X_2+30X_3+480X_4+450X_5+2300X_6+650X_7+300X_8 - X_{17} = 0; \text{ де:}$$

$X_1, X_2, \dots, X_8$  – площі сільськогосподарських культур (коефіцієнти при змінних – затрати на 1 га, грн.);

$X_{17}$  – загальні затрати на виробництво продукції, грн.

24-те обмеження характеризує затрати на виробництво товарної продукції:

$$600X_1+480X_2+30X_3+480X_4+450X_5+2300X_6+650X_7 - X_{18} = 0; \text{ де:}$$

$X_1, X_2, X_3, X_4, X_5, X_6, X_7$  – площі товарних сільськогосподарських культур (коефіцієнти при змінних – затрати на 1 га, грн.);

$X_{18}$  – загальні затрати на виробництво товарної продукції, грн.



### Лекція 4\_3. Кореляція та регресія

В агрономічних дослідженнях часто зустрічаються такі співвідношення між змінними, коли кожному значенню змінної  $x$  відповідає не одне, а множина можливих значень ознаки  $y$ , тобто їх розподіл. Такі зв'язки, що знаходяться лише при масовому вивченні ознаки, на відміну від функціональних називаються *стохастичними* (ймовірнісними) чи *кореляційними*.

При вивченні кореляційних зв'язків виникає два основних питання – про тісноту зв'язків і про форму зв'язків. Для вимірювання тісноти і форми зв'язків використовують спеціальні статистичні методи, які називають кореляцією і регресією.

За формою кореляція може бути лінійною та криволінійною, за напрямом прямою та оберненою. Кореляцію і регресію називають *простою*, якщо досліджується зв'язок між двома ознаками, і *множинною*, коли вивчається залежність між трьома і більше ознаками.

Під *регресією* розуміють зміну результативної ознаки  $y$  (функції) при певній зміні одного чи декількох факторіальних (аргументів).

Зв'язок між функцією і аргументом виражається рівнянням регресії чи кореляційним рівнянням. При простій регресії рівняння позначається  $y = f(x)$  і при множинній  $y = f(x, y, z, \dots)$

Під лінійною (прямолінійною) кореляційною залежністю між двома ознаками  $x$  та  $y$  розуміють таку залежність, яка носить лінійний характер і виражається рівнянням прямої лінії  $y = a + bx$ . Це рівняння називається рівнянням регресії  $y$  на  $x$ , а відповідна йому пряма лінія – вибірковою лінією регресії  $y$  на  $x$ . Коли при однакових приростах аргументу функція має неоднакові зміни, регресія називається криволінійною.

Лінійна регресія  $y$  на  $x$  показує, як змінюється в середньому величина  $y$  при зміні величини  $x$ . Якщо при збільшенні  $x$  величина  $y$  в середньому збільшується, то кореляція і регресія називаються позитивною чи прямою, а якщо зі збільшенням  $x$  значення  $y$  в середньому зменшується – від'ємною чи оберненою.

Для аналізу лінійної регресії між  $x$  та  $y$  проводять  $n$  незалежних парних спостережень, результатом кожного із яких є пара чисел  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . По цим значенням визначають вибіркові емпіричні коефіцієнти кореляції та регресії, розраховують рівняння регресії, будують теоретичну лінію регресії та оцінюють значимість отриманих результатів.

В першу чергу необхідно з'ясувати, яка загальна форма залежності  $y$  від  $x$ . Корисним методичним прийомом є побудова графіка.

Звичайно на графіку вдається візуально знайти передбачувану форму зв'язку. Стандартні програми аналізу регресій, що є практично у всіх обчислювальних центрах (а також в програмно-математичному забезпеченні всіх сучасних персональних комп'ютерів), дозволяють методом перебору підібрати найбільш

відповідну до даного випадку форму зв'язку. На практиці найчастіше дослідження починають з розгляду і оцінки лінійного зв'язку як найпростішої і зручнішої для інтерпретації параметрів.

Форму зв'язку виразимо рівнянням прямої лінії:

$$y = a + bx$$

де  $y$  — значення результативної ознаки;  $x$  — значення ознаки фактора;  $a$  і  $b$  — шукані параметри рівняння.

Якби значення результативного признака  $y$  змінювалися строго пропорційно градаціям факторного признаку  $x$ , то очікувані значення  $y_i$  можна було б точно визначити по значеннях  $x_i$ , що задаються. У реальних умовах спостережувані значення  $x$ , відрізнятимуться від очікуваних значень  $\tilde{y}_i$  на величину  $\varepsilon_i$  тобто є деякий розподіл відхилень:

$$\varepsilon_i = y_i - \tilde{y}_i.$$

Чим менше ця різниця, тим ясніше виражена закономірність зв'язку між ознаками. Тому при визначенні параметрів необхідно відшукати таку форму зв'язку, яка б забезпечила мінімум відхилень. Оскільки відхилення мають різні знаки (розсіяння навколо середньої величини), висувається вимога забезпечення мінімуму суми квадратів відхилень:

$$\varepsilon = \sum_{i=1}^n (y_i - \tilde{y}_i)^2 \rightarrow \min$$

де  $n$  — кількість досліджуваних значень,  $n > 2$ .

Метод, в якому невідомі параметри вибираються, щоб задовольнити вимогам, називають **методом найменших квадратів**.

Якщо для величини  $\varepsilon$  застосувати теорію екстремумів, то отримуємо необхідні умови для визначення невідомих параметрів  $a$  і  $b$ :

$$\frac{\partial \varepsilon}{\partial a} = 0,$$

$$\frac{\partial \varepsilon}{\partial b} = 0.$$

Отримані рівняння називають нормальними рівняннями (кількість нормальних рівнянь має дорівнювати кількості параметрів).

Проведемо визначення параметрів рівняння лінійної залежності  $a$  і  $b$  методом найменших квадратів. Щоб визначити невідомі параметри рівняння  $y = a + bx$  необхідно побудувати систему з двох рівнянь (у загальному випадку число рівнянь рівне числу невідомих параметрів) і розв'язати їх щодо цих невідомих.

Система нормальних рівнянь буде мати такий вигляд:

$$\begin{cases} \sum_{i=1}^n y_i = na + b \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i x_i = a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 \end{cases}$$

$$\sum y = na + b \sum x,$$

де  $n$  — число спостережень;  $\sum y, \sum x$  — суми значень признаков.

Приклад. Маємо дані про врожайність озимої пшениці за різних інтенсивностей освітлення, що вимірюються фотосинтетично активною радіацією (ФАР) в ккал на  $1\text{см}^2$ . Результати спостережень занесено в таблицю.

Таблиця 1. Залежність врожайності озимої пшениці від інтенсивності фотосинтетично активної радіації (ФАР)

№ п/п	ФАР, ккал на $1\text{см}^2$	Урожайність, ц з 1га	Розрахункові величини			Очікувані за рівнянням зв'язку значення урожайності, ц з 1га
			$yx$	$x^2$	$y^2$	
	$x$	$y$	$yx$	$x^2$	$y^2$	$\tilde{y}$
1	15	29	435	225	841	28,69
2	17	34	578	289	1156	33,79
3	18	35	630	324	1225	36,33
4	20	42	840	400	1764	41,43
5	21	45	945	441	2025	43,98
6	23	48	1104	529	2304	49,07
7	26	57	1482	676	3249	56,71
Сума	140	290	6014	2884	12564	290,00

Для вирішення системи на основі емпіричних даних визначають необхідні величини:  $\sum y, \sum x, \sum yx, \sum x^2$ , а потім визначають значення шуканих параметрів.

Треба знайти рівняння зв'язку між врожайністю озимої пшениці та інтенсивністю освітлення. Розв'язавши нормальні рівняння, одержимо:

$$a = -9,52. \quad b = 2,5474.$$

Шукане рівняння лінійного зв'язку має вигляд:

$$\tilde{y} = -9,52 + 2,5474x.$$

Отримане рівняння описує характер зв'язку між ознаками і називається рівнянням регресії.

У багатьох дослідженнях виявляється, що деяка результативна ознака змінюється під впливом не одного, а кількох факторів. Зокрема аналізуючи агрономічні системи на їх роботу впливають дві групи факторів: виробничо-економічні, пов'язані із застосуванням добрив, селекцією і насінництвом, рівнем агротехніки, матеріальною зацікавленістю працівників сільськогосподарського виробництва і т.д. Усі вони цілком залежать від людини. До другої групи факторів належать метеорологічні фактори: опади, температура повітря, висота сніжного покриву, вологість ґрунту і т.д. Ці фактори не залежать від людини і діють хаотично.

Розрізняють лінійні та нелінійні моделі; одно факторні та багатofакторні.

### ***L\_5.***

#### ***Основи аналізу рівнянь регресії***

Рівняння регресії дозволяє оцінити роль факторної ознаки у формуванні результативної. У якості числового показника лінійної кореляції, що вказує на тісноту і напрям зв'язку  $x$  з  $y$ , використовують коефіцієнт кореляції ( $r$ ). Він є безрозмірною величиною, що змінюється в області  $-1 < r < 1$ . Коефіцієнт кореляції знаходиться за формулою:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Чим ближче  $r$  до 1 чи -1, тим тісніший прямолінійний кореляційний зв'язок; він послаблюється з наближенням  $r$  до 0. При  $r = 0$ , між  $x$  та  $y$  немає лінійного зв'язку, але криволінійна залежність може існувати.

Для цього, спираючись на метод розкладання варіації по факторах, необхідно визначити частку фактора в загальній варіації (змінності) результативної ознаки.

Загальна варіація (змінність) врожайності ( $y_i$ ) може бути охарактеризована показником загальної дисперсії:

$$\sigma_{\text{заг}}^2 = \frac{\sum_{i=1}^n y_i^2}{n} - \bar{y}^2, \quad (1)$$

де  $\sigma_{\text{заг}}^2$  - загальна дисперсія врожайності,  $\bar{y}$  - середня врожайність.

В нашому прикладі маємо

$$\sigma_{\text{заг}}^2 = 78,5.$$

Загальна змінність результативного показника  $y$ , що аналізується, обумовлюється двома групами причин:

—впливом змін фактора  $x$ , що включений до моделі;

—впливом множини випадкових причин, що не враховуються в моделі.

Тому постає задача — кількісно виміряти долю впливу включеного до моделі фактора  $x$  та випадкової компоненти.

Оскільки одержане за методом найменших квадратів рівняння  $y = ax + b$  у середньому характеризує закономірний зв'язок між  $x$  та  $y$ , ми можемо вважати, що теоретично очікувані значення  $y$  формуються під «чистим» впливом змін  $x$  та їх варіація з'являється саме за рахунок впливу даного фактора.

Тоді дисперсія результативної ознаки, що обумовлена впливом закладеного до моделі фактора, визначається:

$$\sigma_{\text{фактор}}^2 = \frac{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{n}, \quad (2)$$

де, як було прийнято,  $\bar{y}$  — середнє значення результативної ознаки,  $\tilde{y}_i$ , — теоретично очікувані за рівнянням регресії значення тієї ж ознаки,  $n$  — число спостережень.

Дисперсію, що викликається випадковою компонентою, визначають за відхиленнями теоретичних значень результативної ознаки від фактичних:

$$\sigma_{\text{випадкова}}^2 = \frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{n}, \quad (3)$$

Загальна дисперсія дорівнює сумі факторної та випадкової (залишкової):

$$\frac{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{n} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n} + \frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{n}, \quad (4)$$

тобто

$$\sigma_{\text{загальна}}^2 = \sigma_{\text{фактор}}^2 + \sigma_{\text{випадкова}}^2. \quad (5)$$

Розрахуємо ці величини стосовно до нашого прикладу. Факторна дисперсія буде дорівнювати

$$\sigma_{\text{фактор}}^2 = \frac{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}{n} = 77,9.$$

Залишкова буде дорівнювати

$$\sigma_{\text{випадкова}}^2 = \frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{n} = 0,6.$$

Тоді

$$\sigma_{\zeta\alpha\alpha}^2 = \sigma_{\delta\alpha\delta}^2 + \sigma_{\text{in}\delta}^2 = 77,9 + 0,6 = 78,5.$$

Поділивши праву та ліву частини останньої рівності на загальну дисперсію, матимемо співвідношення:

$$1 = \frac{\sigma_{\delta\alpha\delta}^2}{\sigma_{\zeta\alpha\alpha}^2} + \frac{\sigma_{\text{in}\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}, \quad (6)$$

$$1 = 0,992 + 0,008,$$

де питома вага факторної варіації в загальній

$$\eta^2 = \frac{\sigma_{\delta\alpha\delta}^2}{\sigma_{\zeta\alpha\alpha}^2} = 0,992.$$

Чим більше змінність результату залежить від закладеного в модель фактора, тим ближче значення питомої ваги до одиниці. Звідси випливає, що відношення  $\frac{\sigma_{\delta\alpha\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}$  можна розглядати як критерій впливу фактора на змінність результативної ознаки, як вимірювач тісноти зв'язку між цією та факторною ознаками.

Наступне відношення називають відношенням детермінації:

$$\eta^2 = \frac{\sigma_{\delta\alpha\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}, \quad (7)$$

$$1 = \eta^2 + \frac{\sigma_{\text{in}\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}, \quad \eta^2 = 1 - \frac{\sigma_{\text{in}\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}. \quad (8)$$

де  $\sqrt{\eta^2}$  називають *коефіцієнтом детермінації*. Цей показник є універсальним вимірювачем тісноти зв'язку між ознаками і змінюється він також від 0 до 1.

Стосовно нашого прикладу його значення дорівнює

$$\eta = \sqrt{1 - \frac{\sigma_{\text{in}\delta}^2}{\sigma_{\zeta\alpha\alpha}^2}} = 0,996.$$

За умов повного функціонального зв'язку між ознаками  $\eta = 1$ , якщо зв'язок не знаходиться  $\eta = 0$ .

Універсальність розглянутих показників  $\eta$  та  $\eta^2$  обумовлена тим, що факторна дисперсія, використовувана для їх розрахунку, визначається на основі теоретично очікуваних значень результативної ознаки  $\tilde{\delta}_i$  (за встановленим рівнянням регресії, причому саме рівняння не обов'язково може бути лінійним). Логіка аналізу залишається тією ж і у випадку нелінійної залежності між ознаками.



### Метод введення часу (параметра) в рівняння регресії

У природі не існує явищ, що вільні від динамічності. Будь-який процес, особливо у багатовимірних системах має елементи динамічності. Математичне моделювання таких процесів має ряд проблем. Складність цих проблем часто зумовлює обмеження: вивчення динамічного процесу проводиться за окремими (дискретними) моментами часу. Одержуючи, таким чином, певні моделі для фіксованого моменту часу, слід зробити узагальнення для проміжних моментів та, якщо можливо, для наступних. Як відомо, отримані результати за допомогою простих рівнянь регресії часто не збігаються з дослідними даними. Тому розробка методів одержання рівнянь регресії для опису багатовимірних динамічних систем має велике значення.

Нехай отримано деякі лінійні або нелінійні рівняння регресії. Розглянемо метод побудови динамічної або параметричної багатофакторної моделі на основі одержаних рівнянь.

Нехай на багатовимірній системі, що вивчається, проведено достатню кількість досліджень в кожне дискретне значення часу, та з допомогою деякого методу отримано ряд рівнянь регресії для опису стану багатовимірної системи у відповідні моменти часу: для моменту  $t = t_0$ :

$$y^0 = a_0^0 + \sum_{j=1}^m a_{1j}^0 x_j + \sum_{j=1}^m a_{2j}^0 x_j^2 + \dots + \sum_{j=1}^m a_{kj}^0 x_j^k + \sum_{j=2}^m b_{1j}^0 x_1 x_j + \sum_{j=3}^m b_{2j}^0 x_2 x_j + \dots + b_{m-1m}^0 x_{m-1} x_m + \dots;$$

для моменту часу  $t = t_1$

$$y^1 = a_0^1 + \sum_{j=1}^m a_{1j}^1 x_j + \sum_{j=1}^m a_{2j}^1 x_j^2 + \dots + \sum_{j=1}^m a_{kj}^1 x_j^k + \sum_{j=2}^m b_{1j}^1 x_1 x_j + \sum_{j=3}^m b_{2j}^1 x_2 x_j + \dots + b_{m-1m}^1 x_{m-1} x_m + \dots;$$

.....  
для моменту часу  $t = t_l$

$$y^l = a_0^l + \sum_{j=1}^m a_{1j}^l x_j + \sum_{j=1}^m a_{2j}^l x_j^2 + \dots + \sum_{j=1}^m a_{kj}^l x_j^k + \sum_{j=2}^m b_{1j}^l x_1 x_j + \sum_{j=3}^m b_{2j}^l x_2 x_j + \dots + b_{m-1m}^l x_{m-1} x_m + \dots;$$

де  $a_0, a_{ij}, b_{ij}$  — коефіцієнти рівнянь регресії,  $x_j$  — аргументи.

Для кожного часового моменту визначено значення коефіцієнтів рівнянь регресії. Зрозуміло, що коефіцієнти цих рівнянь будуть відмінні для кожного фіксованого моменту часу. Загалом же кожний коефіцієнт може бути записаний функцією часу. Багатовимірна система, що досліджується, є динамічною, тому її математичний опис за допомогою наближених рівнянь регресії із сталими коефіцієнтами не відповідає

реальності. За допомогою одержаних послідовних наближених рівнянь регресії з сталими коефіцієнтами можна наближено описати лише окремі дискретні стани досліджуваної багатовимірної системи. Для повного ж математичного описання динаміки процесу необхідно провести вивчення коефіцієнтів як невідомих функцій часу.

Для дослідження системи невідомих функцій можна застосувати метод інтерполювання, який виражається в наступному. Під інтерполюванням розуміють пошук значень функції, що відповідають проміжним значенням аргументу. Тобто, мова йде про знаходження аналітичного виразу для многочлену, що приймає в заданих точках задані значення функції. Таким чином, треба побудувати алгебраїчні вирази для невідомих коефіцієнтів основних рівнянь регресії, які були вказані вище. Як аргумент невідомих функцій виступає в даному випадку час:

$$a_0 = f_1(t),$$

$$a_{ij} = f_{ij}(t),$$

$$b_{ij} = \varphi_{ij}(t),$$

.....

Звичайно, в якості алгебраїчних функцій  $f_1(t), f_{ij}(t), \varphi_{ij}(t)$  беруть поліноми аргументу  $t$ . Причому ступінь поліному визначається числом точок часу, для яких побудовано основні регресійні моделі. Це дозволяє проводити функції  $f_1(t), f_{ij}(t), \varphi_{ij}(t)$  чітко через значення коефіцієнтів рівнянь регресії  $a_0, a_{ij}, b_{ij}$ .

В інтервалах між фіксованими моментами часу  $t_1, t_2, \dots$  функції  $f_1(t), f_{ij}(t), \varphi_{ij}(t)$  дозволяють шляхом інтерполяції визначити значення коефіцієнтів рівнянь регресії тієї ж структури, що і одержані раніше рівняння регресії для будь-яких моментів  $t_{11}, t_{12}, \dots, t_{21}, t_{22}, \dots$  при плавній зміні незалежної змінної — часу.

Покажемо, як розв'язується задача введення в рівняння регресії значення часу як параметра на прикладі простих випадків.

Нехай для двох фіксованих часових значень знайдено лінійні рівняння регресії, що зв'язують вихід біосистем  $y$  з одним вхідним впливом  $x$ :

$$\text{для } t = t_1 \quad y = a_{01} + a_{11}x;$$

$$\text{для } t = t_2 \quad y = a_{02} + a_{12}x.;$$

Тут коефіцієнти рівнянь  $a_{01}, a_{11}, a_{02}, a_{12}$  не залежать від часу, а кожне рівняння через фіксовані значення коефіцієнтів може давати значення виходу біосистеми при зміні значення входу. Але при переході від одного моменту часу ( $t_1$ ) до другого ( $t_2$ ) значення коефіцієнтів  $a_0, a_1$  змінюються. Тому для вводу в рівняння регресії часу як незалежної змінної необхідно записати коефіцієнти  $a_0, a_1$  у вигляді функцій часу. Оскільки є тільки два рівняння регресії (для моментів часу  $t_1$  та  $t_2$ ), то як алгебраїчну модель можливо взяти такі лінійні поліноми від часу:

$$a_0(t) = b_0 + b_1 t;$$

$$a_1(t) = c_0 + c_1 t.$$

Тепер два рівняння регресії для моментів часу  $t_1$  та  $t_2$  можна перетворити в одне, що вміщує як незалежні змінні вхідний фактор  $x$  та час  $t$ :

$$y = b_0 + b_1 t + c_0 x + c_1 t x.$$

Відзначимо, що одержане рівняння відповідає нелінійному рівнянню регресії, яке вміщує добуток аргументів. Справді, рівняння можна переписати у вигляді:

$$y = b_0 + b_1 t + c_0 x + c_1 t x.$$

Тепер задача зводиться до визначення коефіцієнтів  $b_0, b_1, c_0, c_1$  рівняння. Їх повинні визначити так, щоб при  $t = t_1$  та  $t = t_2$  рівняння відповідало б раніше добутих рівнянням регресії відносно фактору  $x$ . Для цього застосуємо значення  $a_0, a_1$  а, для фіксованих моментів часу  $t = t_1$ , та  $t = t_2$ . Складемо такі дві системи:

$$\begin{cases} a_0(t_1) = a_{01} = b_0 + b_1 t_1, \\ a_0(t_2) = a_{02} = b_0 + b_1 t_2, \\ a_1(t_1) = a_{11} = c_0 + c_1 t_1, \\ a_1(t_2) = a_{12} = c_0 + c_1 t_2. \end{cases}$$

Тепер розв'яжемо їх стосовно невідомих коефіцієнтів  $b_0, b_1, c_0, c_1$ .  
Отримаємо:

$$b_0 = \frac{a_{01}t_2 - a_{02}t_1}{t_2 - t_1},$$

$$b_1 = \frac{a_{02} - a_{01}}{t_2 - t_1},$$

$$c_0 = \frac{a_{11}t_2 - a_{12}t_1}{t_2 - t_1},$$

$$c_1 = \frac{a_{12} - a_{11}}{t_2 - t_1}.$$

Звідки запишемо вирази для визначення значень коефіцієнтів  $a_0, a_1$ :

$$a_0(t) = \frac{a_{01}t_2 - a_{02}t_1 + (a_{02} - a_{01})t}{t_2 - t_1},$$

$$a_1(t) = \frac{a_{11}t_2 - a_{12}t_1 + (a_{12} - a_{11})t}{t_2 - t_1}.$$

Одержані формули визначають єдиним чином шукані поліноми для визначення значень коефіцієнтів  $a_0, a_1$  стосовно часу як незалежної змінної, в яких значення коефіцієнтів визначені через коефіцієнти початкових факторних рівнянь регресії та фіксовані моменти часу для  $t = t_1$  та для  $t = t_2$ . В закінченому вигляді шукане рівняння регресії набуде вигляду:

$$y = \frac{a_{01}t_2 - a_{02}t_1 + (a_{02} - a_{01})t}{t_2 - t_1} + \frac{a_{11}t_2 - a_{12}t_1 + (a_{12} - a_{11})t}{t_2 - t_1} x.$$

В точках  $t = t_1$  та  $t = t_2$  це рівняння дає два початкових рівняння регресії.

Отримане рівняння регресії застосовують не тільки для моментів  $t = t_1$  та  $t = t_2$ , а й для будь-яких значень  $t$  з інтервалу  $[t_1, t_2]$ .

Аналогічно можуть бути знайдені залежності коефіцієнтів від часу для більш складних рівнянь регресії.

Є ще один спосіб врахування часу в регресійних моделях. Він враховує, що час, як один з факторів в багатфакторному рівнянні регресії, входить до рівняння у вигляді

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

або

$$y = f(t) + f_1(x_1, x_2, \dots, x_n).$$

Але таке часове введення не враховує можливої зміни факторів за часом, яка в попередньому способі враховувалася через зміну коефіцієнтів регресії за часом.

Якщо зміну стану багатовимірної системи обумовлено не змінною часу, а деякою іншою змінною параметра, то аналогічно можливо побудувати математичну модель, що враховує динаміку даного параметра. В цьому випадку математична модель, що описує стан багатовимірної системи, називається параметричною.

## L\_6.

### Поняття про експертні оцінки

Експертні оцінки – це кількісні оцінки процесів або явищ, що не піддаються безпосереднім вимірам. Експертні оцінки мають суттєву вагу при прийнятті планових рішень (при ранжуванні цілей, прогнозуванні альтернатив та їх наслідків). Грунтуються експертні оцінки на твердженнях спеціалістів, що можуть бути індивідуальні або колективні. Поділяються на дві групи.

Перша група методів забезпечує послідовне покращення оцінок кожного експерта. До цієї групи зараховують метод послідовних порівнянь: спочатку експерт визначає попередні оцінки за запропонованою йому шкалою, потім ставляться деякі питання до різних комбінацій результатів, що дозволяють отримати інформацію, на базі якої можливо коригувати попередній набір оцінок.

До другої групи методів зараховують методи: переваги, рангу, часткового попарного порівняння, повного попарного порівняння. Ці методи мають узгоджувати позиції спеціалістів задля вироблення колективної експертної оцінки.

Розглянемо, наприклад, метод рангу. Нехай маємо  $m$  цілей та  $n$  експертів. У кожного з експертів існує своя кількісна шкала переваги цілей у діапазоні 0 - 10. Треба знайти усереднені ваги цілей. Максимальна оцінка надається цілі, що має найбільшу перевагу. Оцінки можуть бути як цілими, так і ні. Позначимо оцінку цілі  $i$  експерта  $j$  через  $p_{ij}$  ( $0 \leq p_{ij} \leq 10$ ). Потім розраховуються відносні ваги кожної цілі для кожного експерта за формулою:

$$v_{ij} = \frac{p_{ij}}{\sum_{i=1}^m p_{ij}}.$$

Після цього обраховують усереднену вагу кожної цілі  $v_i$ , стосовно суджень усіх експертів:

$$v_i = \frac{\sum_{j=1}^n v_{ij}}{\sum_{j=1}^n \sum_{i=1}^m v_{ij}}.$$

Серед методів отримання групової експертної оцінки особливу увагу заслуговує метод Дельфі. За ним опитування експертів (без їх безпосереднього спілкування один з одним) проводиться в декілька турів на основі ретельно розробленої програми. В першому турі кожний експерт відповідає на питання, що поставлені в анкеті. Відповіді усіх учасників обробляються і в наступній анкеті повідомляється про результати опитування. Потім учасників просять переглянути та при бажанні виправити свої попередні відповіді. В цьому турі експертам доводять до відома не тільки отримані результати, а й коротку довідку пояснень, які виставлені на захист думок, що сильно різняться поміж собою. Як правило, опитування за Дельфійським методом обмежується чотирма турами, а медіану кінцевих відповідей приймають за оцінку, що найбільш близька до колективної думки усієї групи.

Характерною особливістю моделювання та прогнозування соціально-економічних процесів є багатоваріантність, тобто можливість використання різних методів, моделей, інформаційного забезпечення, критеріїв оцінювання адекватності моделі тощо. Вибір між конкуруючими варіантами базується на певній системі правил, що забезпечують надання обґрунтованих оцінок кожному варіанту.

Уважається, що експерт (лат. *expertus* — досвідчений) володіє цією системою правил і може порівняти варіанти, приписуючи кожному з них числа. Найчастіше перевага чи відносна значущість варіантів встановлюється за допомогою методів ранжування, попарних порівнянь або безпосереднього оцінювання.

При ранжуванні експерт повинен розмістити варіанти (фактори, моделі, об'єкти тощо) у порядку, який вважає раціональним, і приписати кожному з них числа натурального ряду — ранги 1, 2, ...,  $n$ . Кількість рангів дорівнює кількості варіантів. Якщо експерт надає двом і більше варіантам однакові ранги, то кожному з цих варіантів приписується середній ранг, обчислений з відповідних чисел натурального ряду.

При обґрунтуванні складних управлінських рішень в умовах невизначеності, при довгостроковому прогнозуванні розвитку науки, техніки, економіки використовують групові експертизи. Надійність групових оцінок залежить від узгодженості думок експертів, що потребує відповідної статистичної обробки інформації.

При груповій експертизі ( $n$  експертів) для кожного  $i$ -го варіанта визначається сума рангів  $\sum R_i$ , за якою упорядковуються варіанти. Скажімо, перший — найвищий — ранг надається варіанту, який набирає найменшу суму рангів, а останній — варіанту з найбільшою сумою рангів. Результати опитування експертів оформляються у вигляді матриці.

Наприклад, за даними ранжування трьох варіантів п'ятьма експертами (табл. 1.1), перший ранг надається варіанту А, для якого  $\sum R_i = 6$ , другий — варіанту В, третій — варіанту С. Слід зазначити, що ранги визначають лише місця варіантів поміж іншими, не враховуючи існуючих між ними відстаней.

Таблиця 1.1

Варіант	Експерт					Сума рангів	d	d <sup>2</sup>
	1	2	3	4	5			
<u>A</u>	<u>2</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>6</u>	<u>-4</u>	<u>16</u>
<u>B</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>2</u>	<u>2</u>	<u>10</u>	<u>0</u>	<u>0</u>
<u>C</u>	<u>3</u>	<u>3</u>	<u>2</u>	<u>3</u>	<u>3</u>	<u>14</u>	<u>4</u>	<u>16</u>
<u>Разом</u>	<u>X</u>	<u>X</u>	<u>X</u>	<u>X</u>	<u>X</u>	<u>30</u>	<u>X</u>	<u>32</u>

Статистична обробка результатів ранжування передбачає оцінювання ступеня узгодженості думок експертів. Мірою узгодженості слугує коефіцієнт конкордації  $W$ , в основу розрахунку якого покладено відхилення  $d$  сум рангів за окремими варіантами  $\Sigma R_i$  від середньої суми рангів, яка становить  $\frac{1}{2} n (m + 1)$ . Коефіцієнт конкордації — це відношення суми квадратів названих відхилень  $S = \Sigma d^2$  до максимально можливої суми квадратів відхилень  $S_{\max} = n^2 (m^3 - m) / 12$ . Якщо ранги не повторюються, то

$$W = \frac{12S}{n^2(m^3 - m)},$$

де  $m$  — кількість варіантів;

$n$  — кількість експертів.

При неузгодженості думок експертів  $W = 0$ . Чим вищий ступінь узгодженості, тим більше значення  $W$  наближається до 1. За даними табл. 1.1, середня сума рангів становить  $30 : 3 = 10$ , сума квадратів відхилень  $S = 32$ , а коефіцієнт конкордації

$$W = (12 \cdot 32) / 52 (33 - 3) = 0,64,$$

що свідчить про певні розбіжності в оцінках експертів щодо значущості варіантів.

Перевірка істотності коефіцієнта конкордації  $W$  здійснюється за допомогою критерію  $\chi^2$  з  $(m - 1)$  числом ступенів вільності (свободи). Статистична характеристика критерію розраховується за формулою  $\chi^2 = Wn(m - 1)$ . Для наведеного прикладу  $\chi^2 = 0,64 \times 5(3 - 1) = 6,4$ , що перевищує критичне значення  $\chi^2(2) = 5,99$  (див. додаток 2). Це дає підстави стверджувати з імовірністю 0,95, що значення  $W = 0,64$  не випадкове і думки експертів узгоджені.

При попарних порівняннях експерти використовують дві оцінки: 0 або 1. Більш вагомому варіанту надається оцінка 1, менш вагомому — 0. Результати попарних порівнянь оформляються у вигляді матриці, елементами якої є кількості наданих переваг  $a_{ij}$ . Діагональні елементи такої матриці представлені нулями. Одна із властивостей матриці  $a_{ij} + a_{ji} = n$ , де  $n$  — кількість експертів. За результатами опитування (табл. 1.1) матриця кількості переваг має такий вигляд (табл. 1.2):

Таблиця 1.2

Варіант	A	B	C	Разом	$\omega_i$
A	0	4	5	9	0,60
B	1	0	4	5	0,33
C	0	1	0	1	0,07
Разом	1	5	9	15	1,00

Відношення кількості наданих відповідному варіанту переваг до загальної суми елементів матриці характеризує його вагомість. За даними табл. 1.2, найвагомішим виявився варіант А, для якого  $\omega = 9 : 15 = 0,60$ .

Часто завданням експерта є не ранжування варіантів, а безпосереднє оцінювання рівнів певного явища чи окремих його властивостей, скажімо, якості продукції, конкурентоспроможності фірм тощо. У таких ситуаціях спершу визначається шкала (діапазон) оцінок, у межах якої експерт і оцінює явище (властивість) певним балом  $z_{ij}$ , де  $i$  — властивість,  $j$  — елемент сукупності.

Для певної множини  $m$  властивостей одного явища визначається середній бал  $G_j = \sum z_{ij} / m$ .

На таких методичних засадах ґрунтується більшість рейтингових систем. Так, всесвітньо відома рейтингова система CAMEL, якою користуються органи нагляду за банківською діяльністю, має п'ятибальну шкалу оцінок: від 1 (добре) до 5 (незадовільно). Для кожного банку оцінюється достатність капіталу, якість активів, ефективність менеджменту, прибутковість і ліквідність балансу. Середній бал  $G_j$  є рейтингом фінансового стану  $j$ -го банку. Від його значення залежить ступінь втручання органів банківського нагляду і комплекс заходів щодо усунення недоліків.

Якщо властивості  $z_i$  не рівновагомі, то рейтинг визначається як середня арифметична зважена  $G_j = \sum z_{ij} \omega_i$ , де  $\omega_i$  — вага  $i$ -ї властивості. Саме так оцінюються комерційні, політичні ризики тощо. Наприклад, комерційний ризик, пов'язаний з інтернаціоналізацією банківської діяльності, оцінюється індексом Бері. Ознакова множина цього індексу включає 15 різновагомих показників, які характеризують політичну та економічну ситуацію в країні-партнерові. Зокрема, політична стабільність (вага 12 %), стан платіжного балансу (вага 6 %), темп економічного розвитку (вага 10 %), інші. Сума ваг становить 100 %.

Одним з популярних методів формування групової експертизи є метод Дельфи, назва якого походить від дельфійських мудреців, які славилися в давнину передбаченнями майбутнього. Основні принципи методу Дельфи: анонімність, регульованість зворотного зв'язку та узгодженість групової оцінки.

Автономне опитування експертів проводиться, як правило, в чотири тури. Кожного разу експерт виражає свою думку певною оцінкою в межах визначеної шкали. Результати опитування групи експертів упорядковуються; на основі упорядкованого ряду визначається медіана  $M_e$  й квартилі оцінок — нижній  $Q_1$  і верхній  $Q_3$ . Медіана розглядається як узагальнююча групова оцінка процесу; для характеристики варіації оцінок використовують інтерквартильний розмах  $R = Q_3 - Q_1$ .

Значення медіани і розмаху повідомляють усім експертам. Тим з них, чий оцінки виявилися за межами діапазону ( $Q_3 - Q_1$ ), пропонують аргументувати свої висновки, аби ознайомити з ними решту експертів. Такий зворотний зв'язок відсікає «шуми», зменшує вплив індивідуальних і групових інтересів, не пов'язаних з проблемою.

Ітераційна процедура упорядкування та узагальнення експертних оцінок дає можливість зблизити точки зору експертів, що робить групові оцінки надійнішими



за просте усереднення. Проте сама по собі процедура опитування не розв'язує всіх проблем точності прогнозів. Вирішальну роль відіграють компетентність експертів і досконалість програми опитування.

### Лекція №7-8

#### *Моделі типу «хижак—жертва» без врахування вікової структури*

Втрати врожаю від шкідників значні і мають місце майже завжди. Цим пояснюються зусилля дослідників, що спрямовані на поглиблене розуміння механізму дій визначаючих факторів та на організацію контролю за розмірами популяцій шкідників з тим, щоб економічний збиток від них не перевищував заданої межі.

Класичний біологічний контроль оснований на регулюванні розмірів популяції шкідників за допомогою їх природного ворога. Пізніше з'явилася концепція інтегрованого контролю, що поєднує в собі як біологічні, так і хімічні засоби захисту рослин.

Розглянемо ряд простих динамічних моделей, що описують системи типу «хижак—жертва».

Нехай для даного природного середовища  $y$  позначає число жертв, а  $x$  — число хижаків. Прийнято такі припущення:

1. Діє закон великих чисел; популяція виду може бути представлена однією змінною; вік та стать в розрахунок не беруться.
2. Ефекти миттєві, і немає інтервалів між, наприклад, відкладанням яєць та зрілістю.
3. За відсутності хижаків ( $x=0$ ) число жертв ( $y$ ) росте за логістичною функцією з темпом росту у початковий момент  $\mu\left(\frac{1}{y} \cdot \frac{dy}{dt}\right)$ , і максимальним числом жертв для даного природного середовища  $y_m$ .
4. Темп споживання жертв пропорційний добутку розмірів популяції жертв та хижаків з точністю до сталої  $k$ .
5. Темп виплоду нових хижаків дорівнює  $k', \frac{k'}{k}$  є виплоджування хижаків у перерахунку на 1 жертву.
6. Показник смертності  $m$  приписаний тільки хижакам, для яких жертви не більше як джерело їжі.

Цим припущенням відповідають такі рівняння:

$$\frac{dy}{dt} = \mu y \left(1 - \frac{y}{y_m}\right) - kxy;$$

$$\frac{dx}{dt} = k'xy - mx,$$

де  $\mu, y_m, k, k', m$  — параметри.

Стаціонарні розв'язки, якщо такі існують, можуть бути одержані шляхом зведення  $\frac{dy}{dt}$  та  $\frac{dx}{dt}$  в рівняннях до нуля:

$$0 = \mu \left( 1 - \frac{y}{y_m} \right) - kx; \quad 0 = k'y - m, \quad \text{звідси}$$

$$y = \frac{m}{k'}; \quad x = \frac{\mu}{k} \left( 1 - \frac{m}{k'y_m} \right).$$

Розв'язок буде позитивним тільки у випадку, якщо  $k'y_m > m$ , тобто при умові, що максимальна популяція жертв достатня для перекриття показника смертності хижаків.

Графік розв'язання системи диференціальних рівнянь наближається до спіралі, що закручена за годинниковою стрілкою, ще щоб побудувати цей графік, треба оцінити кількісну динаміку системи (тобто знайти розв'язок рівнянь методами чисельного інтегрування).

Існує багато шляхів, що дозволяють застосованим рівнянням забезпечити більше реалізму або узгодити їх з тими чи іншими спеціальними ситуаціями.

1. Автономний ріст хижаків. Якщо хижаки мають ще інший спосіб споживання, окрім своїх жертв, рівняння приймає вигляд:

$$\frac{dx}{dt} = k'xy + \mu_x x \left( 1 - \frac{x}{x_m} \right),$$

де застосовано припущення, що число хижаків  $x$  зростає за логістичним законом з параметрами  $\mu_x, x_m$ .

2. Обмежене споживання жертв хижаками. Згідно з диференціальним рівнянням, споживання жертв в перерахунку на одного хижака зростає лінійно із зростанням числа  $y$  жертв без усяких обмежень. Це навряд чи відповідає дійсності, і тому правильніше було б рівняння подати у вигляді:

$$\frac{dx}{dt} = k'x \left( \frac{y}{1+ay} \right) - mx,$$

де  $a$  - стала, що є межею  $\frac{k'x}{a}$ , до якої для великих значень  $y$  прагне перший член суми в правій частині.

3. Укриття для жертв. В окремих природних середовищах обмежені групи жертв можуть знаходити укриття, що оберігають їх від хижаків. В цьому випадку рівняння набувають вигляду:

$$\frac{dy}{dt} = \mu y \left( 1 - \frac{y}{y_m} \right) - kx(y - y_0)H(y - y_0);$$

$$\frac{dx}{dt} = k'x(y - y_0)H(y - y_0) - mx,$$

де  $y_0$  — максимальне число жертв, що спроможні використати укриття;  $H$  — функція Хевісайда:

$$H(y - y_0) = 1 \text{ при } y - y_0 \geq 0, H(y - y_0) = 0 \text{ при } y - y_0 < 0.$$

Стационарні розв'язки та динамічна поведінка цих модифікованих рівнянь подібні до вище розглянутих.

Розглянемо наявність запізнень в системі. Загальний для усіх запізнень ефект — це внесення додаткової нестабільності. Запізнені реакції в біології можуть мати різне походження. Розглянемо такі випадки.

1. Час, що потрібен для еволюційних явищ. Яка-небудь зміна в довкіллі (наприклад, збільшення ресурсів) може викликати раптовий ріст продуктивності дорослих особин. Однак відповідна зміна чисельності дорослих особин відбудеться тільки через деякий час  $\tau$ , що дорівнює часу, необхідному для розвитку дорослої особини з яйця. Стосовно збільшення біологічної популяції  $y$  іноді приймають

припущення про експоненційну залежність  $\frac{dy}{dt} = \mu y(t)$ , де  $\mu$  — темп росту.

Таким чином, швидкість збільшення  $y$  залежить від чисельного значення  $y$  та часу  $t$ . Так якщо, наприклад, між яйцекладкою та формуванням дорослої особини проходить час  $\tau$ , то правильніше було б записати

$$\frac{dy}{dt} = \mu y(t - \tau),$$

де  $y(t - \tau)$ , — «доросла» популяція в момент  $t - \tau$ .

2. Дискретність сезонів розмноження. Багато видів розмножуються тільки у визначений проміжок року. Навіть в таких популяціях, особини яких здатні розмножуватися декілька років підряд (багаторічні рослини, птахи тощо), наявність сезонів розмноження вносить деяке запізнення в процеси регуляції чисельності. Якщо життєвий цикл певного виду продовжується декілька років та його особини щорічно дають відносно невелику кількість потомства, то запізнення на один рік, що обумовлене дискретністю сезонів розмноження, не вважається довгим порівняно з характерним часом динаміки цього виду. Таким чином, будь-які коливання, що зумовлені запізненням, будуть затухаючими. Якщо ж дорослі особини, що розмножуються в даному році, рідко або ніколи не доживають до розмноження в наступному (наприклад, однорічні рослини, більшість комах тощо), тоді наявний суттєвий вплив на динаміку їх чисельності. В цьому випадку застосовують кінцево-різницеve рівняння вигляду

$$y_{n+1} = \Phi(y_n),$$

де  $y_n$  - чисельність популяції в  $n$ -му році,  $\Phi$  — деяка функція) яке часто приводить до коливань, що розходяться.

3. Впливи зовнішнього середовища, що запізнюються, які обмежують чисельність. Параметр темпу росту  $\mu$  в експоненційній залежності може залежати від змінних середовища  $E(t)$ , до числа яких може входити і ресурс поживного середовища. Впливи цих факторів на організм іноді проявляються із запізненням, і тому більш адекватною була б формалізація

$$\frac{dy}{dt} = \mu[E(t - \tau)]y(t),$$

в якій значення  $\mu$  обчислюється з величин змінних середовища  $E$  з урахуванням затримки реакції на час  $\tau$ .

Наприклад, обмеження, що пов'язані з доступом їжі, можуть давати малий вплив на темп росту популяції, якщо термін дії процесу достатньо великий.

Розглянемо такий приклад. Нехай  $Y$  є розміром популяції, і споживання якої підтримується на постійному рівні  $f$ ; потрібність дорослих особин популяції в їжі складає  $mY$ , де  $m$  — стала; надлишки їжі  $f - mY$  — перетворюються в яйця з ефективністю  $\eta$ , так що темп яйцекладіння складає  $\eta(f - mY)$ ; яйця перетворюються в дорослу особину через час  $\tau$ , смертність дорослих особин має постійний показник  $h$ .

З врахуванням цих умов маємо диференціальне рівняння:

$$\frac{dY(t)}{dt} = \eta[f - mY(t - \tau)] - hY(t).$$

Стаціонарний розв'язок  $Y_s$  при цьому шукають у формі  $Y(t) = Y(t - \tau) = Y_s$ . В результаті отримаємо  $Y_s = \frac{\eta f}{\eta m + h}$ .

Інша система, що природним шляхом приводить до затримки реакцій, — це епідемія з інкубаційним періодом  $\tau$ . Нехай  $Y_s$  — загальна популяція,  $Y$  — уражена,  $(Y_m - Y)$  — популяція, що знаходиться під впливом захворювання. Враховуючи затримку, викликану наявністю інкубаційного періоду, ймовірність ураження залежить від розміру інфікованої популяції в момент  $t - \tau$ , що можна виразити рівнянням

$$\frac{dY(t)}{dt} = \mu Y(t - \tau) \left( 1 - \frac{Y(t)}{Y_m} \right) - cY(t),$$

де  $\mu$  — ймовірність ураження особини в одиницю часу,  $c$  — темп природного одужання.

Стаціонарний розв'язок одержують у формі  $Y_s = Y_m \left( 1 - \frac{c}{\mu} \right)$ . Після заміни

$Y = Y_s + y$  рівняння можна представити у вигляді

$$\frac{dy}{dt} = -\mu y(t) + c y(t - \tau).$$

## Моделі типу «хижак—жертва» з врахуванням вікової структури

В динаміці популяції простіший підхід до моделювання опирається на оцінку повного числа організмів  $N$ . Однак моделі такого типу є надмірно спрощеними, і тому неточними.

Функцію видовікового розподілу  $n(t, a)$  визначають так, що  $n(t, a)da$  — це число організмів, вік яких лежить в межах від  $a$  до  $da + a$  в момент  $t$ . Розмір популяції тоді обчислюється як

$$N(t) = \int_0^{\infty} n(t, a) da.$$

Функції народження та смертності  $B(t, a), D(t, a)$  можна визначити як показники цих процесів в момент  $t$  для організмів віку  $a$ , так що

$$\text{Народжуваність} = \int_0^{\infty} B(t, a)n(t, a)da,$$

$$\text{Смертність} = \int_0^{\infty} D(t, a)n(t, a)da.$$

Організм народжується у віці  $a = 0$ , і тому  $n(t, 0) = \text{народжуваності}$ .

Група організмів віку  $a$  в момент  $t$  втрачає особин внаслідок смерті (насамперед в результаті старіння). Після закінчення періоду часу  $\Delta t$  за відсутності смертності число особин  $n(t, a)da$  збільшується до  $n(t + \Delta t, a + \Delta a)$ , і (показати самостійно) це приводить до рівняння

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = 0,$$

а при наявності смертності

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial a} = -D(t, a)n(t, a).$$

Ще більшої адекватності моделі можна досягнути, якщо застосувати функцію розподілу «вік — розмір — вид», де  $n(t, a, \omega)dad\omega$  позначає число особин у віці від  $a$  до  $da + a$  з розмірами від  $\omega$  до  $\omega + d\omega$ . Загальна кількість особин в популяції тоді визначається як

$$N = \int_0^{\infty} \int_0^{\infty} n(t, a, \omega)dad\omega.$$

Такий метод швидко приводить до дуже складних формалізацій, що сильно ускладнює розв'язання задачі.

Розглянемо метод, який називають **методом матриць Леслі**. На рис. 1 неперервний видовіковий розподіл розбито на 4 когорти. Застосовуючи позначення рисунка, можна побудувати матрицю переходів

$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix}^{(t=i+1)} = \begin{pmatrix} B_1 & B_2 & B_3 & B_4 \\ S_{12} & 0 & 0 & 0 \\ 0 & S_{23} & 0 & 0 \\ 0 & 0 & S_{34} & 0 \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix}^{(t=i)}$$

Коефіцієнти  $B$  позначають народжуваність у 4 когортах,  $S$  - це показники виживання. Наприклад,  $S_{12}$  позначає фракцію першої когорти, яка після закінчення одиниці часу виживає та переходить до складу другої когорти. Матричне рівняння еквівалентне наступним:

$$n_1(i+1) = B_1 n_1(i) + B_2 n_2(i) + B_3 n_3(i) + B_4 n_4(i);$$

$$n_2(i+1) = S_{12} n_1(i);$$

$$n_3(i+1) = S_{23} n_2(i);$$

$$n_4(i+1) = S_{34} n_3(i).$$

Такий підхід широко застосовується в практичних дослідженнях.

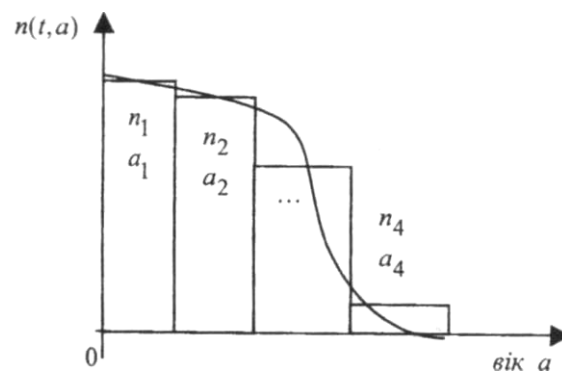


Рис. 1. Функція видовікового розподілу популяції, що зображена кривою, ап-роксимованою чотирма квантилями  $n_1, n_2, n_3, n_4$ .