

# Апроксимація експериментальних даних кривою Безьє

Т.М.Романенко, В.В.Вишневський, В.Г.Калмиков

Інститут проблем математичних машин і систем  
romanenko@immsp.kiev.ua

## Abstract

Very often the experimental data is the realization of the process, fully determined by some unknown function, distorted by hindrances. Treatment and experimental data analysis is substantially facilitated, if to represent these information as analytical expression. The algorithm of the experimental data processing on the example of oncologic researches of preparations of blood is represented in this article.

## 1. Вступ

Дуже часто експериментальні дані є спотвореною перешкодами реалізацією певного процесу, цілком визначуваного деякою невідомою функцією. Оброблення і аналіз експериментальних даних, представлених набором вимірювань, наприклад, у вигляді графіка, таблиці, істотно полегшується, якщо вдається перетворити ці дані до аналітичного виразу. У цій роботі розглядається алгоритм оброблення експериментальних даних на прикладі результатів дослідження препарату крові за методом "Онкотест" [1].

Первинні дані про дослідження препарату крові використовуються для постановки діагнозу про наявність або відсутність онкологічного захворювання у обстежуваного пацієнта. Метод раннього діагностування злоякісних новоутворень "Онкотест" базується на визначенні в крові людини специфічних речовин – маркерів пухлинного росту, що з'являються в період виникнення і розвитку злоякісної пухлини [2]. Для визначення наявності в крові маркерів пухлинного росту автори методу досліджують препарати спеціально обробленої сироватки крові засобами порівняльної спектрофотометрії. Визначають порівняльну оптичну щільність досліджуваного і базового препаратів в діапазоні довжин хвиль 360 - 500 нм по формулі:

$$D = Lg(I/I_0), \quad (Б) \quad (1.1)$$

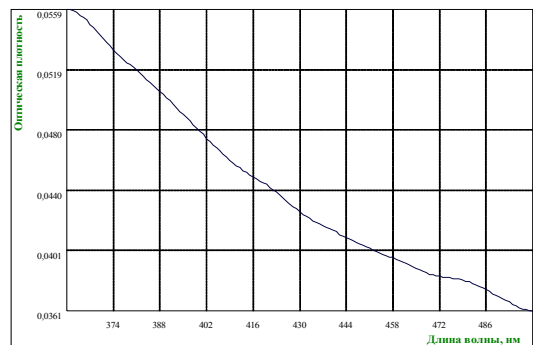


Рисунок 1.1: Спектрограма здорової людини

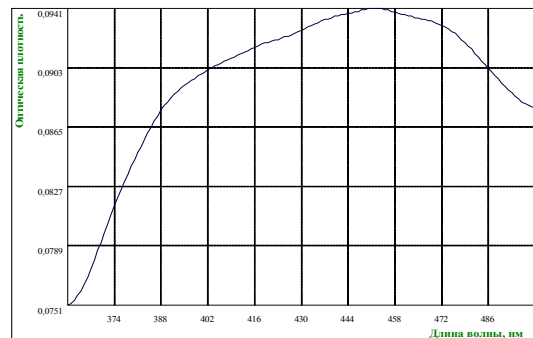


Рисунок 1.2: Спектрограма хворої людини

де  $D$  – порівняльна оптична щільність досліджуваного препарату,

$I$  – інтенсивність світла, що пропускається, в базовому препараті,

$I_0$  – інтенсивність світла, що пропускається, в досліджуваному препараті.

Результатом первинного дослідження препарату крові є спектрограма – залежність оптичної щільності від довжини хвилі у вигляді

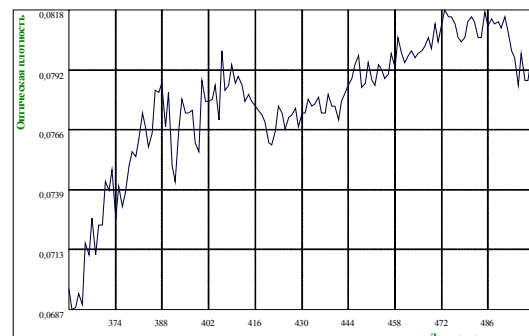
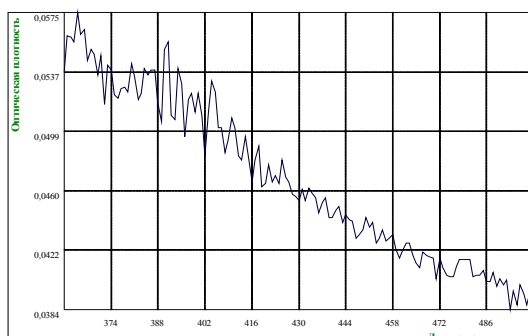


Рисунок 1.3: Експериментальні дані, спотворені перешкодами

графіка. Вид спектрограми людей, в організмі яких розвивається злоякісна пухлина, відрізняється від спектрограми здорових людей. У здорових людей крива оптичної щільності зі збільшенням довжини хвилі, зазвичай, убуває монотонно (рис. 1.1), а у хворих, як правило, монотонність порушується – мають місце екстремуми (рис. 1.2). Даний ефект успішно використовується для ранньої діагностики онкологічного захворювання. Доцільно припустити, що розрізняється також вигляд кривих для окремих локалізацій захворювання. Вигляд спектрограми препарату, одержаного за методом "Онкотест" залежить від наявності або відсутності в сироватці крові маркерів – певної кількості молекул спеціальної речовини, тобто кожен графік є реалізацією деякого образу. Можна вважати, що образом в даному випадку є дуга деякої кривої. У найпростішому випадку мають на увазі тільки два образи, що відповідають: 1 – позитивному і 2 – негативному діагнозу. Для негативного діагнозу характерна крива, функція якої  $D = f(\lambda)$  монотонно убуває зі зростанням  $\lambda$  (рис.1.1). Для позитивного діагнозу характерна наявність, принаймні, одного екстремуму - максимуму функції  $D = f(\lambda)$ . Безпосереднє використання статистичних методів розпізнавання для автоматизованої постановки діагнозу в даному випадку не є можливим, через велику мінливість значень оптичної щільності (десятки разів), та, навіть, конфігурації функції для реалізацій одного й того ж самого образу. Тому вбачається доцільним для кожного графіка знайти деяку найбільш відповідну криву, коефіцієнти функції якої, після приведення до єдиного масштабу, можуть бути використані як параметри образу, інваріантні до змін оптичної щільності. Завдання розділення множини діагнозів на два, а, можливо, і більше класів як статистичне завдання розпізнавання доцільно вирішувати в просторі параметрів деякої кривої, заданої аналітично. Визначення параметрів аналітично заданої кривої, що відображає спектрограму, і є метою даної роботи.

Додаткові ускладнення надає те, що експериментальні дані найчастіше спотворені перешкодами (рис. 1.3).

Таким чином, вважаємо, що науковий і практичний інтерес має представлення даних у вигляді, зручнішому для подальшої обробки та автоматизації діагностики.

## 2. Формальна постановка задачі

Спектрограма може бути представлена ламаною лінією, точки зламу якої утворені послідовністю значень оптичної щільності  $D(n) = \{d_0, d_1, d_2, \dots, d_n, \dots, d_N\}$ , де  $n = 1, N$  – кількість вимірювань оптичної щільності в графіку.

З великого числа часто вживаних кривих найбільш відповідним вбачається використання кривої Безьє.

Під кривою Безьє розуміємо ділянку плоскої кривої третього порядку, задану в параметричному

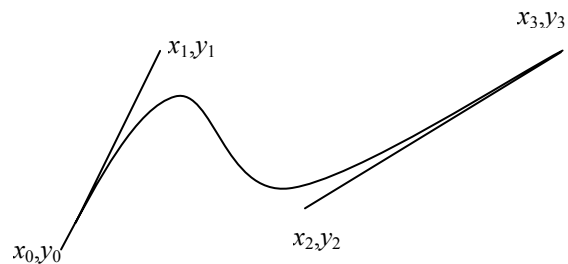


Рисунок. 2.1: Крива Безьє

вигляді таким чином [3]:

$$Bz(t) = \sum_{j=0}^3 Br_j(t) \cdot Q_j, \quad (2.1)$$

де  $Br_j(t)$  – базові поліноми Бернштейна третього ступеня,

$Q_j$  – коефіцієнти кривої,

$t$  – дійсний параметр, що змінюється в інтервалі  $[0,1]$ .

Поліноми Бернштейна визначаються за наступними формулами:

$$Br_j(t) = C_3^j \cdot (1-t)^{3-j} t^j, \quad 0 \leq j \leq 3, \quad (2.2)$$

де  $C_3^j$ ,  $0 \leq j \leq 3$  – біноміальні коефіцієнти,

$$C_3^j = \frac{3!}{j!(3-j)!} \quad (2.3)$$

Таким чином, крива Безьє (рис. 2.1) може бути задана двома поліномами третього ступеня, залежними від параметра  $t$ :

$$X(t) = x_0 \cdot (1-t)^3 + 3 \cdot x_1 \cdot (1-t)^2 \cdot t + 3 \cdot x_2 \cdot (1-t) \cdot t^2 + x_3 \cdot t^3 \quad (2.4)$$

$$Y(t) = y_0 \cdot (1-t)^3 + 3 \cdot y_1 \cdot (1-t)^2 \cdot t + 3 \cdot y_2 \cdot (1-t) \cdot t^2 + y_3 \cdot t^3$$

Представлення експериментальних даних взагалі, і, зокрема, тих, що визначають препарати "Онкотест", у вигляді кривої Безьє має ряд переваг.

Описуючи достатньо широкий клас кривих, криві Безьє використовують відносно невелику кількість коефіцієнтів – 8. Важливою перевагою з погляду користувача-діагноста є наочність. Коефіцієнти  $(x_0, y_0)$ ,  $(x_3, y_3)$  є координатами початкової і кінцевої точок дуги апроксимуючої кривої, відповідно, а коефіцієнти  $(x_1, y_1)$ ,  $(x_2, y_2)$  – координатами точок, що управляють її формою (так званих «вусів»). Змінюючи коефіцієнти або переміщаючи відповідні їм точки на екрані монітора (як початкову і кінцеву, так і ті, що управляють), користувач може в інтерактивному режимі редагувати форму апроксимуючої кривої, покращуючи результати автоматичної апроксимації, а разом з тим і діагностичний графік може бути представлений вектором  $v = \{x_0, y_0, x_1, y_1, x_2, y_2, x_3, y_3\}$ , компонентами якого є коефіцієнти рівнянь кривої Безьє. Опис спектрограми у вигляді вектора  $v$ , після його нормування, може бути використаний в статистичній постановці завдання розпізнавання діагнозу. Визначення коефіцієнтів кривої Безьє, що найбільш точно відтворює ту або іншу спектрограму, не є тривіальним завданням. У роботі [4] для визначення коефіцієнтів кривої Безьє, що відображає послідовність експериментальних точок, використаний ітераційний спосіб мінімізації суми квадратів відстаней від кожної точки графіка до апроксимуючої кривої. Припускається, що точки графіка є взаємно незалежними. Для обчислення відстані від кожної точки графіка до апроксимуючої кривої повинна бути знайдена відповідна найближча точка на апроксимуючій кривій, що саме по собі при параметричному завданні кривої є ітераційним завданням. При значній кількості точок на графіку і наперед не певній кількості ітерацій такий спосіб розв'язання задачі привів до істотних обчислювальних труднощів.

У даній роботі пропонується мірою схожості деякої апроксимуючої кривої і даного експериментального графіка прийняти площу  $S(v, D(n))$  фігури, обмеженої контуром, утвореним згаданими кривою, визначуваною коефіцієнтами  $v$ , і графіком  $D(n)$ . Тоді пошук кривої Безьє з коефіцієнтами  $v_{opt}$ , що найкращим чином апроксимує заданий експериментальний графік, можна розглядати як визначення коефіцієнтів кривої Безьє, що мінімізують значення  $S(v, D(n))$ :

### 3. Алгоритм

$$v_{opt} = \underset{v}{\operatorname{argmin}} S(v, D(n)) \quad (2.5)$$

Як видно з формули 2.5, основою алгоритму пошуку оптимального значення коефіцієнтів  $v_{opt}$  є обчислення площі, обмеженої контуром  $S(v, D(n))$ . Безпосереднє обчислення  $S(v, D(n))$  не вбачається можливим через те, що даний контур не є однозв'язним. Графік  $D_N$  і крива Безьє з коефіцієнтами  $v$  неодноразово перетинаються. Обчислення координат точок перетину утруднене параметричним завданням кривої Безьє. Найбільш простим рішенням представляється заміна кривої Безьє ламаною лінією  $V_t$ , точки зламу якої  $(x_t, y_t)$  належать кривій Безьє і обчислені для значень  $t = 0, \delta, 2\delta, 3\delta, \dots, 1$ . Значення  $\delta$  вибрано достатньо малим, щоб забезпечити достатньо точне для практичних обчислень відображення ламаною лінією  $V_t$  початкової кривої.

Результатом прийнятого спрощення є можливість обчислити для кожного значення  $D(n)$  відповідне значення  $V_t(n)$ . Тоді площа  $S(v, D(n))$  може бути обчислена як

$$S(v, D(n)) = \sum_{n=0}^{n=N} |V_t(n) - D(n)| \quad (3.1)$$

використання методу градієнтного спуску.

### 4. Експериментальна перевірка алгоритму

Описаний алгоритм, заснований на методі градієнтного спуску, був експериментально перевірений більш ніж на 1100 реальних спектрограмах (як з позитивним діагнозом, так і з негативним). Результат досягався в більшості випадків за 15 - 40 ітерацій (у особливо складних випадках - до 300). Приклади спектрограм і кривих Безьє, що їх апроксимують, приведені на рис. 4.1 і рис. 4.2. Як видно з рисунків, апроксимуючі криві Безьє дають задовільне наближення.

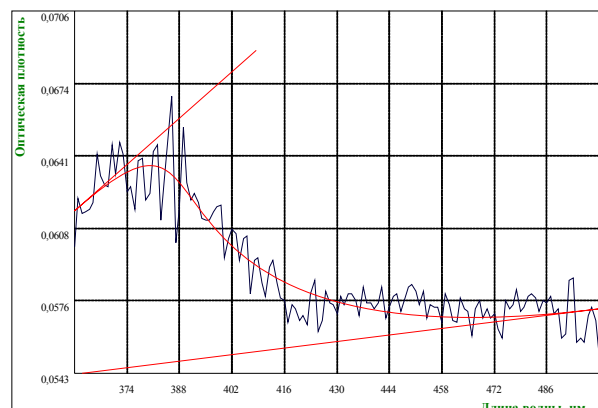
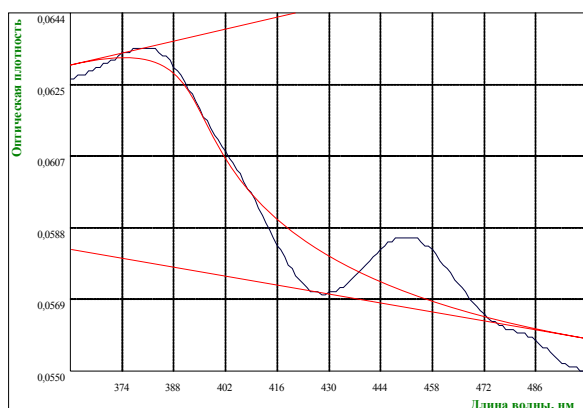


Рисунок 4.1: Приклади апроксимації експериментальних даних спектрограм кривими Безьє (позитивний результат тесту)

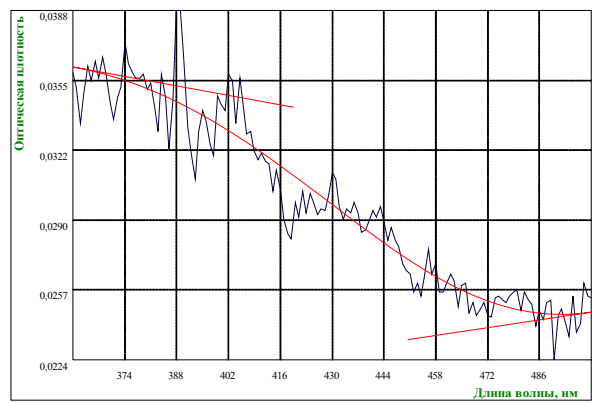
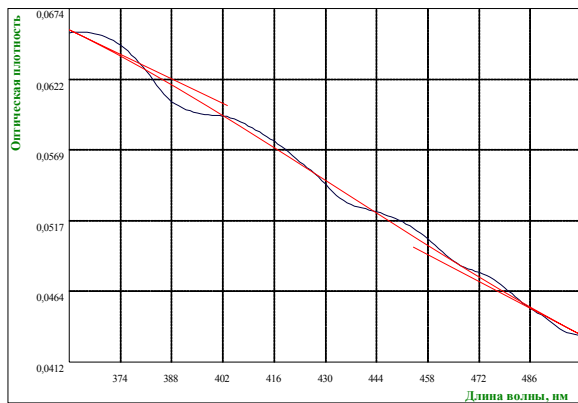


Рисунок. 4.2: Приклади апроксимації експериментальних даних спектрограм кривими Безьє (негативний результат тесту)

## 5. Висновки

Запропонований і реалізований алгоритм дозволяє знайти параметри кривої Безьє, що апроксимує експериментальний графік – спектрограму – з метою виконання автоматизованої діагностики. Це дає можливість автоматизувати процес обстеження населення з метою раннього виявлення злоякісних пухлин.

## 6. Література

- [1] Вишне夫斯基 В.В., Владимиров В.А., Романенко Т.Н. *Программно-аппаратный ком-*

*плекс "Онкотест-WM-01" // Тезиси доповіді II Міжнародної школи-семінару «Телемедицина: досвід та перспективи».*

- [2] *Методичні рекомендації по використанню методу "Онкотест-WM-01" для скринінгу злоякісних пухлин.*
- [3] Денискин Ю.И. *Особенности аппроксимации обводоов параметрическими полиномами в форме Бернштейна // Прикладная геометрия.* – 1999. – вып. 2, № 2
- [4] Вишне夫斯基 В.В., Рисцов І.К., Волжева М.В *Итерационный алгоритм побудови кривої Безьє за заданими точками // Математичні машини і системи.* – 2004, № 4.