

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
KHOA CÔNG NGHỆ THÔNG TIN



TOÁN ỨNG DỤNG THỐNG KÊ

BÁO CÁO THỰC HÀNH

LAB 04

Mã số sinh viên: 21120582

Họ và Tên: Đinh Hoàng Trung.

Mail: 21120582@student.hcmus.edu.vn.

MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

1. Câu 1: Xác định tính lồi lõm của một hàm số:

$$f(x) = x^T A x + b^T x + c$$

- Ý tưởng:
 - Xác định lồi lõm qua giá trị riêng của ma trận A.
 - Nếu hàm lồi hoặc lõm: thì tìm cực tiểu/đại qua việc tìm nghiệm bình phương tối thiểu.
- Thực hiện:
 - Tìm các giá trị riêng qua hàm.
 - Nếu tất cả giá trị riêng đều ≥ 0 , hàm xác định nửa dương \Rightarrow hàm xác định lồi.
 - Nếu tất cả giá trị riêng đều > 0 , hàm xác định dương \Rightarrow hàm xác định lồi ngặt.
 - Nếu tất cả giá trị riêng đều ≤ 0 , hàm xác định nửa âm \Rightarrow hàm xác định lõm.
 - Nếu tất cả giá trị riêng đều < 0 , hàm xác định âm \Rightarrow hàm xác định lõm ngặt.
 - Sau khi xác định lồi/lõm tìm điểm dừng bằng phương pháp bình phương tối thiểu.

2. Câu 2: Tìm đường tuyến tính khớp với dữ liệu cho trước:

- Ý tưởng thực hiện:
 - Phương trình tuyến tính có dạng: $y = ax + b$.
 - Trong đó ta có dữ liệu thực x và y , thứ ta cần tìm là a và b .
 - Tìm a và b bằng phương trình:

$$\sum \begin{bmatrix} 1 \\ x_i \end{bmatrix} [1 \ x_i] \begin{bmatrix} a \\ b \end{bmatrix} = \sum y_i \begin{bmatrix} 1 \\ x_i \end{bmatrix}$$

- Có a và b ta tìm được phương trình tuyến tính khớp với dữ liệu.

MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

- Thực hiện bài toán.

○ Mô hình hóa dữ liệu từ 2 cột giá trị x_i, y_i :

▪ Ma trận với mỗi một dòng là một mẫu $[x_i, y_i]$:

```
[[0, 10],  
 [1, 8],  
 [2, 7],  
 [3, 5],  
 [4, 2]]
```

▪ Ma trận A với mỗi dòng là $[1, x_i]$ và ma trận b với một dòng là 1 mẫu $[y_i]$.

```
[[0. 1.] [10.]  
 [1. 1.] [ 8.]  
 [2. 1.] [ 7.]  
 [3. 1.] [ 5.]  
 [4. 1.] [ 2.]]
```

○ Từ A và b tìm a, b theo phương pháp bình phương tối thiểu

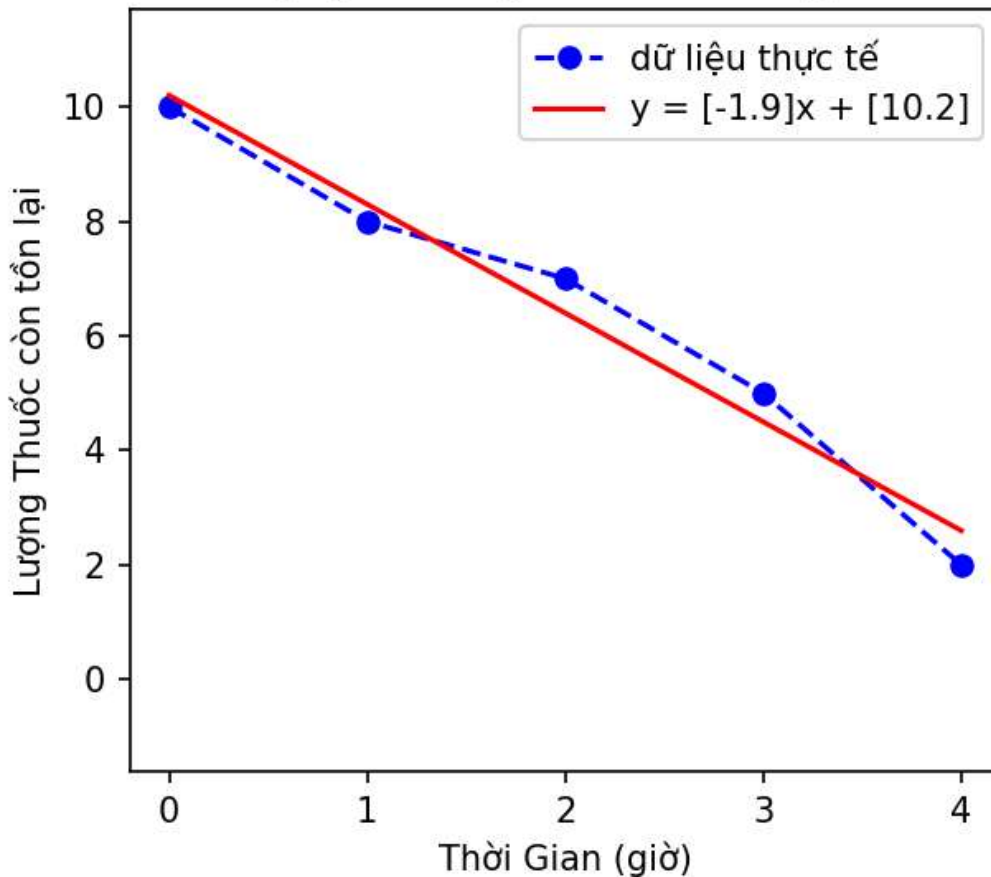
với $A^T A$ là $\sum \begin{bmatrix} 1 \\ x_i \end{bmatrix} [1 \ x_i]$ và $A^T b$ là $\sum y_i \begin{bmatrix} 1 \\ x_i \end{bmatrix}$.

\Rightarrow Từ đó t có phương trình $y = ax + b$.

MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

Đường tuyến tính thể hiện
lượng thuốc giảm theo thời gian:



3. Câu 3:

a. Giả sử x, y thỏa mãn mô hình $y = a + bx + c \cdot \ln(x^2 + 1)$. Hãy sử dụng phương pháp bình phương cực tiểu, ước lượng các tham số a, b, c .

- Ý tưởng làm bài: tương tự bài 2, chỉ đổi biến cần tìm từ a, b sang a, b, c và x, y mô hình hóa theo mô hình trên.
- Thực hiện bài toán:
 - Tương tự như bài 2: Mô hình hóa dữ liệu thực sang mô hình:
 - Từ dữ liệu thực:

MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

$[-2, -1]$,
 $[0, 1.5]$,
 $[1, 3.1]$,
 $[2, 6.3]$,
 $[4, 11.1]$

- Sang dữ liệu được mô hình hóa:

$\begin{bmatrix} 1. & -2. & 1.60943791 \end{bmatrix}$	$\begin{bmatrix} -1. \end{bmatrix}$
$\begin{bmatrix} 1. & 0. & 0. \end{bmatrix}$	$\begin{bmatrix} 1.5 \end{bmatrix}$
$\begin{bmatrix} 1. & 1. & 0.69314718 \end{bmatrix}$	$\begin{bmatrix} 3.1 \end{bmatrix}$
$\begin{bmatrix} 1. & 2. & 1.60943791 \end{bmatrix}$	$\begin{bmatrix} 6.3 \end{bmatrix}$
$\begin{bmatrix} 1. & 4. & 2.83321334 \end{bmatrix}$	$\begin{bmatrix} 11.1 \end{bmatrix}$

- Từ đó theo phương pháp bình phương cực tiểu ta được phương trình khớp với dữ liệu:

Phương trình theo mô hình:

- $y = [1.14446483] + [1.81151861]x + [0.92214453]\ln(x^2 + 1)$

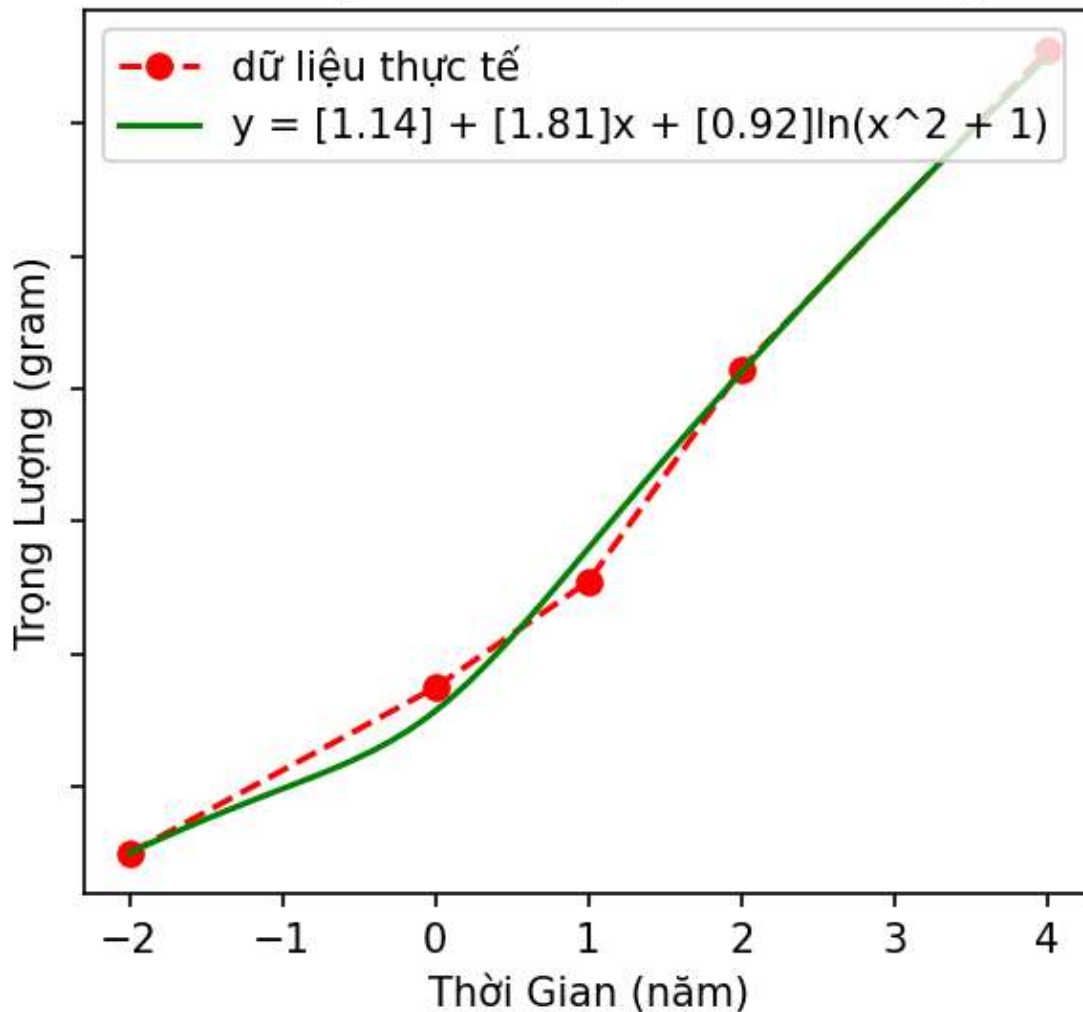
- b. Vẽ biểu đồ và dự đoán y với $x = 6.5$.**

[Với $x = 6.5$, y được dự đoán: $[16.39305168]$ gram.]

MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

Trọng lượng của hợp chất
theo thời gian khi tiếp xúc với không khí



MSSV: 21120582

Họ Tên: Đinh Hoàng Trung.

C. Có nên dùng mô hình $y=a+bx+c\ln(x)$ hoặc $y = a + bx + c/x$ để xấp xỉ dữ liệu trên không.

○ Không vì trong $\ln(x)$ chỉ có nghĩa khi $x > 0 \forall x \in \mathbb{R}$.

Thử mô hình $y = a + bx + c\ln(x)$

```
d:\Toán ứng dụng thống kê\Thực Hành\Lab4\Lab4.py:211: RuntimeWarning: invalid value encountered in log
  A3c[i][2] = np.log(data3[i][0])
d:\Toán ứng dụng thống kê\Thực Hành\Lab4\Lab4.py:211: RuntimeWarning: divide by zero encountered in log
  A3c[i][2] = np.log(data3[i][0])
```

○ Tương tự $1/x$ chỉ có nghĩa khi $x \neq 0 \forall x \in \mathbb{R}$.

Thử mô hình $y = a + bx + c/x$

```
d:\Toán ứng dụng thống kê\Thực Hành\Lab4\Lab4.py:227: RuntimeWarning: divide by zero encountered in scalar divide
  A3c[i][2] = 1/data3[i][0]
```

---HẾT---