**BÀI TẬP PHÒNG HỒI PHỤC**

**MATLAB**

**Mục tiêu:**

Các nhà khoa học và kỹ sư thường cố gắng hiểu một cách định lượng các thí nghiệm và hệ thống mà họ nghiên cứu để thiết kế các giải pháp cho các vấn đề khác nhau hoặc để hiểu các cơ chế cơ bản. Kỹ năng quan trọng để phân tích dữ liệu định lượng bao gồm việc khớp các mô hình với dữ liệu để hiểu liệu một hiện tượng cơ bản có thể giải thích một kết quả thí nghiệm cụ thể hay không. Trong phòng thí nghiệm này, bạn sẽ học cách khớp các mô hình toán học (phương trình) với dữ liệu.

Đặc biệt, bạn sẽ học cách:

1. Điều chỉnh đường cong cho dữ liệu và xác định mức độ phù hợp
2. Sử dụng hàm fminsearch trong MATLAB để thu nhỏ hàm
3. Hiểu từ vựng dùng để mô tả mô hình phù hợp với dữ liệu
4. Sử dụng lý thuyết đơn giản về việc điều chỉnh mô hình để chọn mô hình tốt nhất cho tập dữ liệu

**Các định nghĩa:**

**Đa thức:**Đa thức có dạng

*f(x) = a n x n + a n-1 x n-1 + ... + a 1 x + a 0*Phương trình (1)

Trong đó **n** là **thứ tự** của đa thức.

Ví dụ: f(x) = 2x 2 - 4x + 10 ( Bậc 2)

f(x) = 9 (Bậc 0)

Trong các định nghĩa sau đây, “phương trình” có thể là một đa thức nhưng cũng có thể là bất kỳ phương trình nào khác.

**Thông số**

Giá trị của các hằng số trong phương trình. Ví dụ: [ *a n , a n-1 , …a 1 , a 0* ] là các tham số của đa thức bậc n, [2, -4, 10] là các tham số của ví dụ bậc 2 ở trên và [ C, τ] là các tham số của phương trình f(x) = C\*exp(-x/  ). Khi tham số là hằng số nhân các số hạng (chẳng hạn như “C” chứ không phải “  ” trong ví dụ trước), chúng ta cũng gọi chúng là **hệ số** .

**Lắp đường cong:**

Lắp một phương trình vào một tập hợp các điểm dữ liệu. Điều này được sử dụng để kiểm tra mức độ phù hợp của mô hình định lượng với dữ liệu. Nghĩa là, nếu bạn có một tập dữ liệu ( **x** , **y** ), trong đó **x** và **y** là vectơ, thì khớp đường cong là tìm một phương trình f( **x** ), sao cho f(x(i)) gần với y(i) cho tất cả các chỉ số i trong vectơ.

**Dư:**

Đối với mỗi điểm dữ liệu, phần dư của nó là chênh lệch giữa điểm dữ liệu thực tế và giá trị của phương trình được sử dụng để mô hình hóa dữ liệu. Tức là phần dư là f(x(i)) –y(i).

**Hình vuông nhỏ nhất phù hợp:**

Các tham số của một phương trình cho giá trị thấp nhất của tổng bình phương của tất cả các phần dư. Nghĩa là, phương pháp khớp bình phương nhỏ nhất tìm thấy các giá trị tham số **P** làm cực tiểu hóa hàm sumerror = sum(f( **P** ,x(i)) –y(i)), trong đó f( **P** ,x(i)) là phương trình với tham số P tính tại điểm x(i). Ví dụ: nếu phương trình là đa thức bậc 2 và P là [2,-4,10], thì f( P **,** x(i)) = 2(x(i)) 2 - 4x(i) + 10.

**Động lực cho phòng thí nghiệm:**

Hầu hết các loại thuốc đều có lợi ích điều trị khi thuốc ở trên “ngưỡng điều trị” nhưng lại có tác dụng phụ độc hại trên “ngưỡng độc”. Phác đồ định lượng nhằm mục đích duy trì nồng độ thuốc trong máu bệnh nhân trong một “cửa sổ trị liệu” nằm giữa hai ngưỡng này. Tuy nhiên, chúng ta có thể mở rộng đáng kể thời gian điều trị nếu chúng ta có thể giảm tác dụng phụ bằng cách chỉ nhắm mục tiêu thuốc đến các tế bào cần chúng. Nghiên cứu hiện tại về phân phối thuốc kết hợp nghiên cứu dược phẩm với công nghệ sinh học để kết hợp cả thuốc và phân tử nhắm mục tiêu vào hạt nano, là hạt có đường kính nhỏ hơn 1 micron (thường là vài chục hoặc hàng trăm nanomet). Phân tử nhắm mục tiêu liên kết với thụ thể được ưu tiên biểu hiện trên các tế bào đích trong cơ quan đích, do đó tập trung các hạt nano vào các tế bào này. Các hạt nano cũng cần có cơ chế giải phóng thuốc sau khi được liên kết hoặc hấp thụ bởi các tế bào này.

Để hiểu toàn bộ quá trình phân phối, nhà nghiên cứu có thể tối ưu hóa các quy trình khác nhau liên quan đến việc phân phối thuốc bằng cách sử dụng một số thử nghiệm *in vitro* và *in vivo* . Những điều này có thể bao gồm tối đa hóa khả năng liên kết của các hạt nano với các tế bào đích, giảm thiểu sự liên kết với các tế bào đối chứng (không phải mục tiêu) và giảm thiểu tốc độ thanh thải của các hạt nano ở động vật. Cũng có thể mong muốn tối ưu hóa tốc độ giải phóng thuốc dựa trên các quy trình khác. Nghĩa là, hạt nano phải liên kết với các tế bào mục tiêu của nó trước khi phần lớn thuốc được giải phóng và sau đó giải phóng thuốc một cách hiệu quả. Việc sử dụng các mô hình toán học cho từng quy trình này một cách riêng lẻ và cùng nhau có thể giúp kỹ sư tối ưu hóa thiết kế hạt nano để toàn bộ quá trình phân phối đạt hiệu quả cao nhất có thể và nồng độ thuốc cục bộ được giữ trong cửa sổ trị liệu. Trong bài tập hôm nay, bạn sẽ kiểm tra các mô hình toán học chỉ cho một phần của quá trình này - một thí nghiệm *trong ống* nghiệm về việc giải phóng thuốc từ hạt nano. (“Hệ thống nano cho hình ảnh đồng thời và phân phối thuốc đến tế bào T”, bởi Fahmy và cộng sự, Tạp chí AAPS, 2007 tập 9.)

**Bài tập thí nghiệm:**

Fahmy và cộng sự đã đo tổng lượng thuốc được giải phóng từ hạt nano theo thời gian. Một số kết quả của họ được cung cấp trong bảng dưới đây:

Thuốc Day Microgram được phát hành

0 2.5

0,25 3,6

0,5 5,3

1 9,5

2 14.0

3 16,5

4 18,8

5 21,5

6 23,2

8 26,8

10 28,4

*Lúc đầu, chúng tôi không tìm cách giải thích lý do tại sao dữ liệu lại hoạt động như vậy. Thay vào đó, chúng tôi chỉ đơn giản tìm cách mô tả hành vi đó bằng toán học. Kiểu mô hình hóa này được gọi là mô hình hóa “* ***hộp đen*** *”, với ý tưởng là bạn không thể nhìn xuyên qua hộp mô hình để hiểu quy trình cơ bản. Nó còn được gọi là mô hình “* ***thực nghiệm*** *”, nghĩa là nó bắt nguồn từ quan sát trực tiếp (dữ liệu) chứ không phải bằng lý thuyết về cơ chế cơ bản (ví dụ, lý thuyết về cách thuốc được giải phóng và do đó phương trình toán học nào sẽ mô tả điều đó). quá trình.) Miễn là mô hình thực nghiệm mô tả tốt dữ liệu, nó có thể được sử dụng để dự đoán điều gì sẽ xảy ra khi quá trình này được tích hợp vào một hệ thống lớn hơn.*

1. Viết hàm lấy dữ liệu (ở bất kỳ dạng nào bạn muốn) làm đầu vào, khớp dữ liệu theo một đường thẳng bằng cách sử dụng polyfit và polyval , đồng thời trả về các tham số và giá trị mô hình tại các điểm dữ liệu. Vẽ cả mô hình và dữ liệu trên cùng một biểu đồ. **Bao gồm biểu đồ này và các thông số trong phòng thí nghiệm của bạn viết lên.**
2. Sửa đổi hàm được tạo ở câu hỏi 1 để tính và trả về phần dư. Vẽ phần dư của mô hình dưới dạng biểu đồ thanh. **Đưa âm mưu này vào bài viết trong phòng thí nghiệm của bạn. (** Gợi ý: Kiểm tra lệnh “bar”)
3. Tính tổng bình phương của tất cả các phần dư. **Bao gồm giá trị tính toán trong phòng thí nghiệm của bạn viết lên.**
4. Sửa đổi chương trình chính và/hoặc hàm để phù hợp với dữ liệu bằng cách sử dụng đa thức **bậc 2** và **bậc 3 . Bao gồm các đồ thị, số dư và tổng bình phương của số dư trong bài viết của bạn** . Lưu ý: nếu bạn đã biết cách sử dụng điều khiển luồng, bạn có thể tạo một vòng lặp để khớp dữ liệu đã cho với đa thức bậc 1 , 2 và 3 , vẽ đồ thị dữ liệu thô và mô hình cho từng dữ liệu đó, đồng thời lưu trữ phần dư cho mỗi mô hình.
5. Thứ tự của đa thức phù hợp nhất về tổng bình phương phần dư thấp nhất là gì?

Dữ liệu đã cho cũng có thể được mô hình hóa bằng phương trình **hàm mũ sau** :

Phương trình này mô hình hóa một tình huống trong đó một hệ thống đi từ giá trị 0 đến giá trị tối đa với hằng số tốc độ . Lần này, các thông số và có ý nghĩa khái niệm (lượng thuốc tối đa được giải phóng và hằng số tốc độ giải phóng) và bắt nguồn từ một mô hình thực sự tìm cách mô hình hóa các quy trình cơ bản. Do đó, mô hình này được gọi là mô **hình hộp xám** (trong suốt một phần) hoặc mô hình **tham số** (các tham số có ý nghĩa vật lý).

1. Sử dụng MATLAB , khớp dữ liệu tương tự vào mô hình mới này, sử dụng các ước tính ban đầu . Vì phương trình này không phải là hàm đa thức nên bạn không thể sử dụng polyfit mà thay vào đó sẽ sử dụng hàm giảm thiểu tổng quát hơn có tên là fminsearch (xem gợi ý ở cuối). Vẽ biểu đồ dự đoán mô hình tốt nhất và dữ liệu trên cùng một biểu đồ. **Bao gồm âm mưu này trong phòng thí nghiệm của bạn viết lên** . **Các thông số phù hợp nhất là gì?**

Có một nguyên tắc gọi là dao cạo Occam, trong đó nói rằng lời giải thích đơn giản nhất giải thích một hiện tượng có thể là lời giải thích đúng. Trong mô hình hóa, điều này có nghĩa là phương trình có ít tham số nhất phù hợp nhất với dữ liệu được coi là mô hình tốt nhất. Có một số lý thuyết định lượng mức độ phù hợp tốt hơn để biện minh cho một tham số bổ sung nằm ngoài phạm vi của bài tập này.

1. **Trả lời câu hỏi sau:** Bạn nên từ chối sử dụng dao cạo của Occam trong bốn mô hình nào và mô hình nào là lựa chọn hợp lệ? Nếu số lượng tham số quan trọng hơn mức độ phù hợp nhất thì mô hình nào là tốt nhất? Nếu sự phù hợp nhất là quan trọng nhất thì mô hình nào là tốt nhất?

Một trong những điều có giá trị nhất mà một mô hình có thể làm là dự đoán hành vi của hệ thống trong những điều kiện ban đầu chưa được đo lường. Chúng tôi đã giữ bạn cho đến bây giờ, chỉ cung cấp cho bạn dữ liệu trong 10 ngày. Tuy nhiên, dữ liệu được công bố trải qua 21 ngày:

Microgam ngày

12 28,4

16 28,5

21 29,5

1. Mô hình của bạn có thể được sử dụng để dự đoán điều gì sẽ xảy ra trong khoảng thời gian từ ngày 10 đến ngày 21: *Không cần điều chỉnh lại mô hình* , hãy tính giá trị của 4 mô hình theo ngày và thêm giá trị này vào biểu đồ của bạn cùng với dữ liệu hoàn chỉnh. **Hãy đưa những biểu đồ này vào bài viết của bạn và trả lời những câu sau:** Mô hình nào có khả năng dự đoán tốt nhất trong 21 ngày? Kết luận này liên quan thế nào đến phân tích của bạn ở phần 7?

**Gợi ý QUAN TRỌNG PHẢI ĐỌC: (** Sau bài thực hành này,bạn có trách nhiệm **hiểu** cách sử dụng fminsearch giúp bạn điều chỉnh đường cong. **)**

Triển khai fminsearch trong MATLAB:

Công việc fminsearch thực hiện: Bắt đầu với những dự đoán ban đầu của bạn cho các tham số (bạn chuyển những tham số này làm đầu vào), fminsearch gọi một hàm sử dụng các tham số đó (bạn đặt tên của hàm này làm đầu vào). Hàm này thường được gọi là “ **hàm mục tiêu** ” và phải tạo ra một giá trị vô hướng duy nhất làm đầu ra (được gọi là giá trị của hàm mục tiêu). Nói chung, bạn sẽ tự viết hàm mục tiêu cho bài toán tối thiểu hóa hoặc tối ưu hóa cụ thể của mình. Trong mỗi vòng lặp, hàm fminsearch thay đổi các tham số một chút và xác định xem giá trị của hàm mục tiêu tăng hay giảm và giữ tham số mới trong trường hợp giảm. Việc lặp lại (vòng lặp) này tiếp tục cho đến khi nó xác định rằng nó đã tìm thấy điểm tối thiểu (cục bộ) cho hàm mục tiêu, sau đó trả về các giá trị tham số cuối cùng (và giá trị cuối cùng của hàm mục tiêu, nếu bạn muốn).

Để sử dụng fminsearch để khớp một phương trình với dữ liệu, bạn viết một hàm mục tiêu tính toán phương trình bạn đang cố gắng khớp, sử dụng các tham số mà nó được truyền cho mỗi điểm (ví dụ: mỗi điểm thời gian) trong tập dữ liệu của bạn. Sau đó, nó so sánh các giá trị mô hình này với các giá trị dữ liệu thử nghiệm để lấy phần dư, tính tổng bình phương của phần dư và trả về giá trị này. Do đó, hàm mục tiêu chỉ đơn giản là tổng bình phương của phần dư giữa mô hình và dữ liệu. Có ý nghĩa, bởi vì đây là những gì bạn muốn giảm thiểu.

Vì bạn cần so sánh mọi thứ với dữ liệu trong rất nhiều hàm, nên sẽ thuận tiện khi lưu trữ nó dưới dạng biến ' **toàn cục** ', có nghĩa là tất cả các hàm của bạn đều có thể nhìn thấy dữ liệu đó mà không cần bạn phải chuyển nó cho chúng. Bạn có thể thực hiện điều này vì không có chức năng nào của bạn thay đổi giá trị của dữ liệu thực tế này, do đó không có cơ hội gây nhầm lẫn.