**Họ và Tên: Nguyễn Vũ Trường Giang**

**MSSV: 20182471**

**2.1. Visualizing the Data**

Trực quan hóa dữ liệu thường cung cấp cái nhìn sâu sắc có giá trị về vấn đề, nhưng thường bị bỏ qua như một phần của quá trình học máy. Chúng ta sẽ bắt đầu bằng cách trực quan hóa tập dữ liệu đơn biến bằng cách sử dụng phân tán 2D âm mưu của đầu ra so với đầu vào. Tuy nhiên, trong lớp này, chúng ta thường sẽ xử lý đa chiều các tập dữ liệu. Một khi chúng ta vượt ra ngoài hai chiều, việc hình dung trở nên khó khăn hơn nhiều. Trong những trường hợp như vậy, chúng tôi phải trực quan hóa từng thứ nguyên riêng biệt hoặc sử dụng các kỹ thuật giảm kích thước (chẳng hạn như PCA) để giảm số lượng tính năng. Sau đó trong khóa học, chúng ta sẽ thảo luận về các kỹ thuật như vậy.

Bạn có thể tải dữ liệu đơn biến vào các biến ma trận X và y bằng cách thực hiện các lệnh sau trong trình thông dịch python từ trong thư mục hw1:

Graphical user interface, text, application

Description automatically generated

Sau đó, tạo biểu đồ phân tán dữ liệu bằng phương thức plotData1D:



Đây là đầu ra kết quả của đoạn code trên:

Chart, scatter chart

Description automatically generated

**2.2. Implementation**

Thực hiện hồi quy tuyến tính đa biến thông qua gradient descent bằng cách hoàn thành khung lớp LinearRegression. Hãy chắc chắn không thay đổi API lớp. Những nơi duy nhất bạn cần thay đổi trong tệp được đánh dấu bằng thẻ nhận xét “TODO”.

Hồi quy tuyến tính phù hợp với vectơ tham số θ với tập dữ liệu. Trong bài tập này, chúng ta sẽ sử dụng phép suy giảm độ dốc để tìm ra giải pháp tối ưu. Nhớ lại rằng hàm mục tiêu hồi quy tuyến tính L2 là hàm lồi, do đó, đường xuống dốc sẽ tìm thấy giá trị tối ưu toàn cục. Vấn đề này cũng có một giải pháp dạng đóng, nhưng nhiều hơn về sau.

Hàm cost function có dạng:

Text

Description automatically generated

Với giả thiết:



Hàm chi phí được viết dưới dạng ma trận có dạng:



Code trong chương trình:

A screenshot of a computer

Description automatically generated

**Gradient Descent**

Gradien Descent được tính theo công thức:



Code trong chương trình:

A screenshot of a computer

Description automatically generated

**Closed Form**

Công thức tính Close Form dưới dạng ma trận:



Code trong chương trình

A screenshot of a computer

Description automatically generated

**Đây là kết quả:**

Chart

Description automatically generated