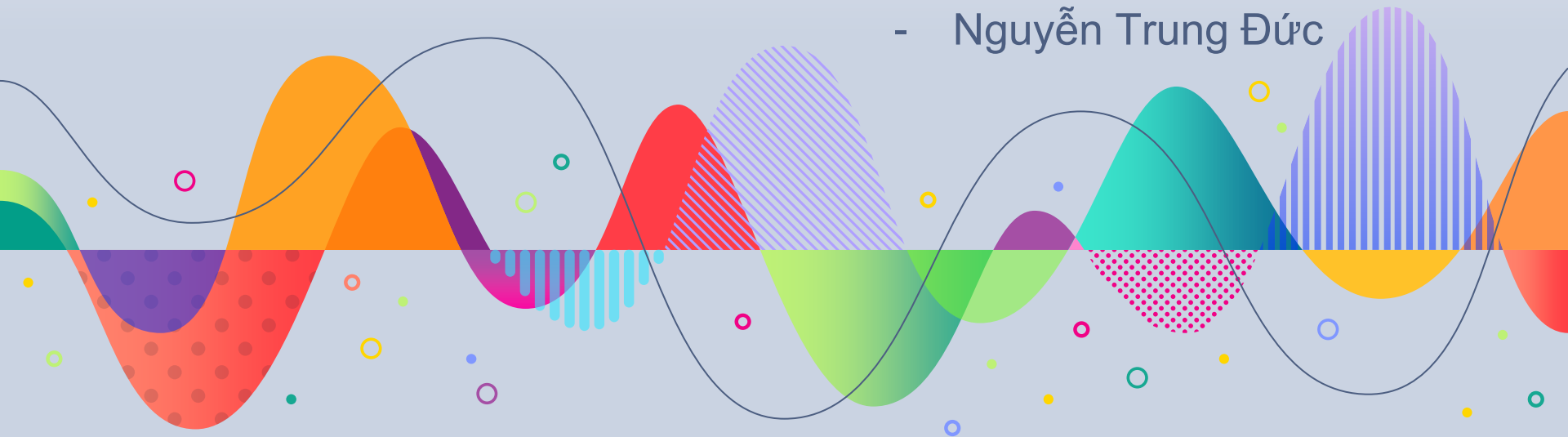


Áp dụng thuật toán Gradient Descent và các biến thể cho dự đoán giá BĐS

- Vũ Minh Hưng
- Nguyễn Trung Đức



1. Giới thiệu dataset



Bộ dữ liệu kc_house về dữ liệu giá trị bất động sản.

- Bộ dữ liệu gồm **21597 bản ghi** với **15 biến đầu vào** là các thông tin của BĐS và biến mục tiêu là giá BĐS
- Dữ liệu đã được lược bỏ bớt và chuẩn hóa

2. Hàm mất mát



Ta sử dụng hàm MSE để tính bình phương chênh lệch giữa giá trị thực tế và giá trị dự đoán, với công thức

$$Cost = \frac{1}{N} \sum_{i=1}^n (y_i - (mx_i + b))^2$$

3. Gradient Descent



Thay vì tìm điểm có đạo hàm bằng 0 (thường rất phức tạp hoặc bất khả thi), hướng tiếp cận phổ biến nhất là:

1. Xuất phát từ một điểm mà ta coi là gần với nghiệm của bài toán
2. Dùng một phép toán lặp để tiến dần đến điểm đạo hàm gần với 0.

Với hàm $f(x_k)$ cần tìm cực tiểu, với đạo hàm tại một điểm x bất kỳ là $\nabla f(x_k)$, với t là step size, đến vòng lặp k , ta cập nhật:

$$x_k = x_{k-1} - t_k \nabla f(x_k),$$

4. Accelerated Proximal method



Với Proximal gradient descent có dạng:

$$x_k = \text{prox}_{t_k h}(x_{k-1} - t_k \cdot \nabla g(x_{k-1})), k = 1, 2, 3, \dots$$

Dựa trên proximal mapping của hàm lồi h :

$$\text{prox}_h(x) = \underset{z}{\operatorname{argmin}} \frac{1}{2} \|x - z\|_2^2 + h(z)$$

4. Accelerated Proximal method



Với $\min g(x) + h(x)$ trong đó g là hàm lồi khả vi và h lồi

Phương pháp Accelerated proximal gradient:

$$x_k = \text{prox}_{t_k h}(v - t_k \nabla g(v)),$$

Phương pháp Accelerated proximal gradient có bước $k = 1$ tương tự như trong proximal gradient. Ở các bước tiếp theo, mỗi bước có v thể hiện 'quán tính' chịu ảnh hưởng bởi các bước trước đó:

$$v = x_{k-1} + \frac{k-2}{k+1}(x_{k-1} - x_{k-2})$$

5. Newton Method



Với hàm tối ưu không ràng buộc, tìm min $f(x)$, với f lồi, khả vi bậc hai, $\text{dom}(f) = \mathbb{R}^2$, phương pháp Newton có công thức lặp:

$$x_k = x_{k-1} - (\nabla^2 f(x_{k-1}))^{-1} \nabla f(x_{k-1})$$

Với $\nabla^2 f(x_{k-1})$ là ma trận Hessian của f tại x_{k-1}

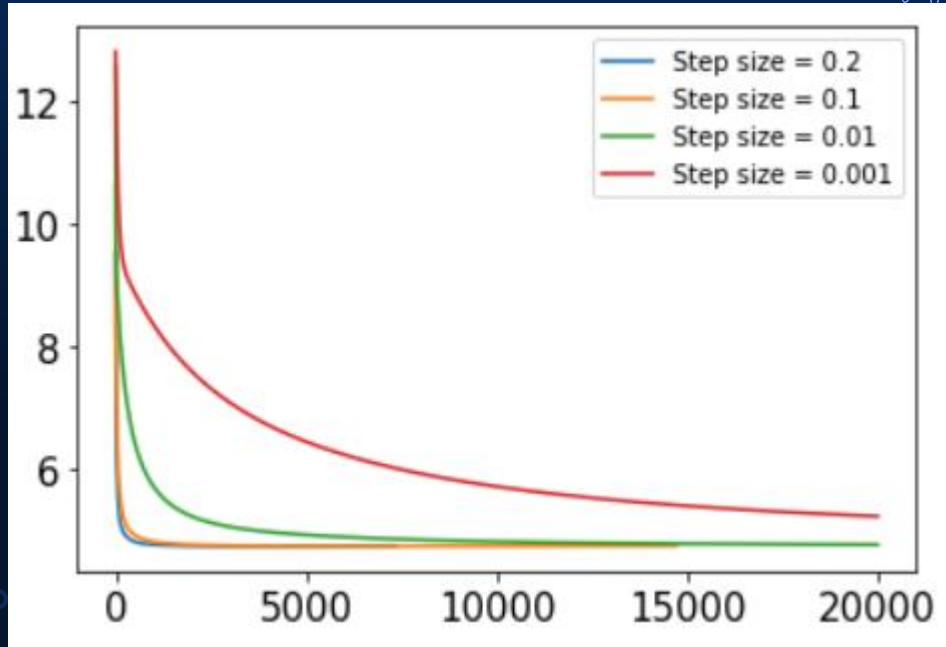
6. Điều kiện dừng



Một vài phương pháp xác định điểm dừng cho thuật toán:

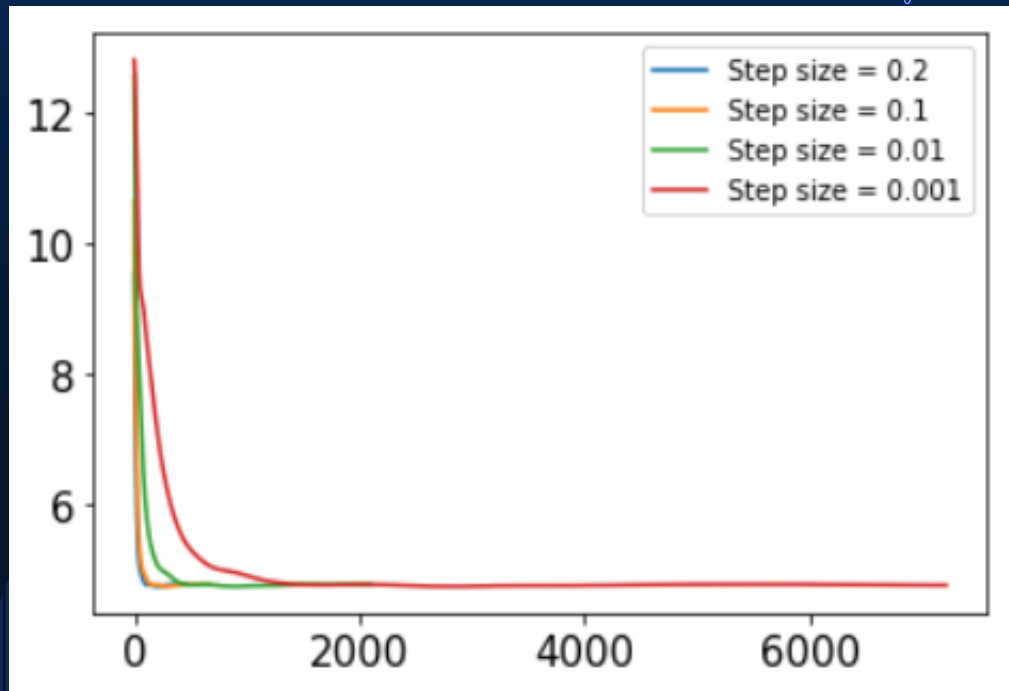
- ▶ **Giới hạn số vòng lặp**
 - ▶ Ưu điểm: đơn giản, phổ biến + đảm bảo thuật toán chạy không quá lâu
 - ▶ Nhược điểm: Thuật toán có thể dừng trước khi đủ gần với nghiệm
- ▶ **So sánh gradient** của nghiệm tại 2 lần cập nhật liên tiếp
 - ▶ Nhược điểm: việc tính đạo hàm có thể trở nên phức tạp

7. So sánh các step size - GD



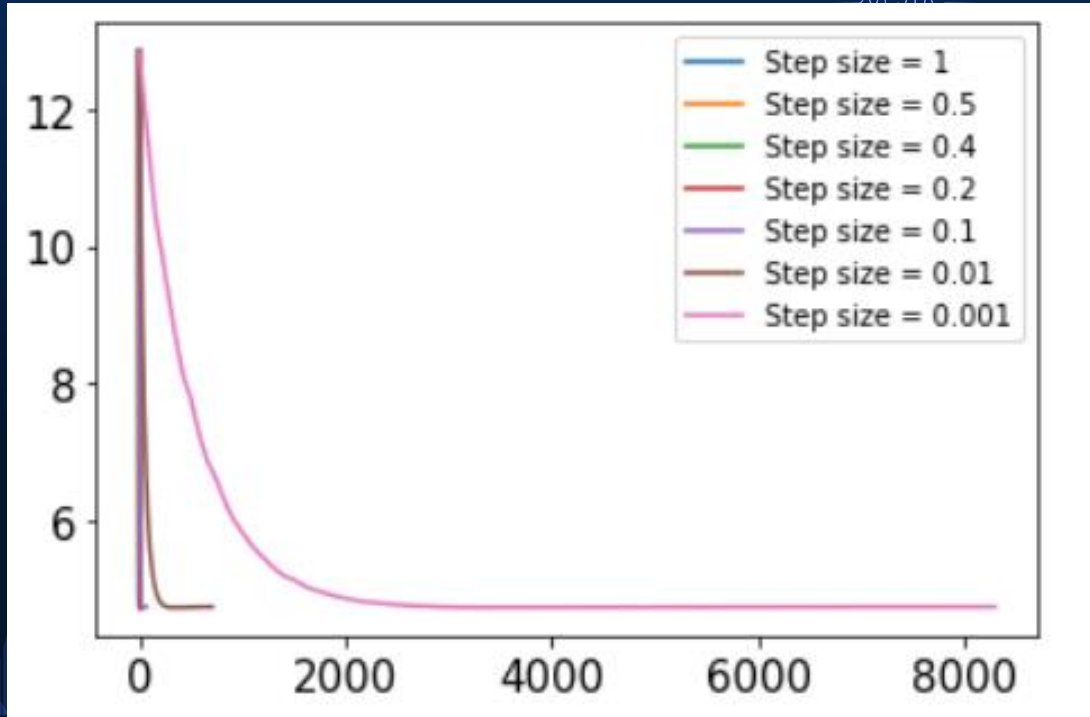
Total steps with size 0.2: 7360
Total steps with size 0.1: 14721
Total steps with size 0.01: 20000
Total steps with size 0.001: 20000

7. So sánh các step size - Accelerated Proximal



Total steps with size 0.2: 466
Total steps with size 0.1: 661
Total steps with size 0.01: 2097
Total steps with size 0.001: 7207

7. So sánh các step size - Newton



Total steps of size 1: 7
Total steps of size 0.5: 12
Total steps of size 0.4: 15
Total steps of size 0.2: 29
Total steps of size 0.1: 68
Total steps of size 0.01: 710
Total steps of size 0.001: 8292

8. So sánh ứng dụng backtracking với các thuật toán



Fitting time of GD: 1.8819658756256104

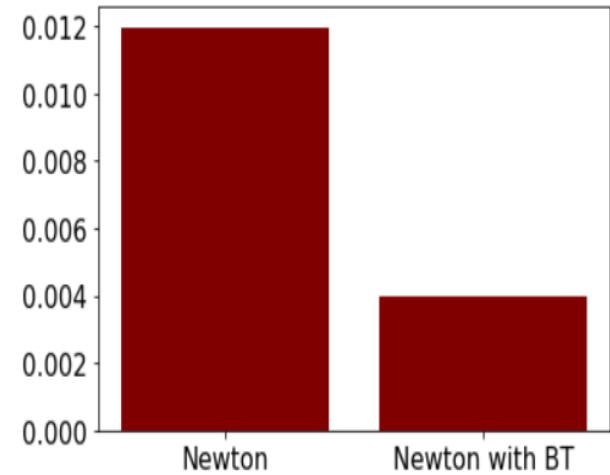
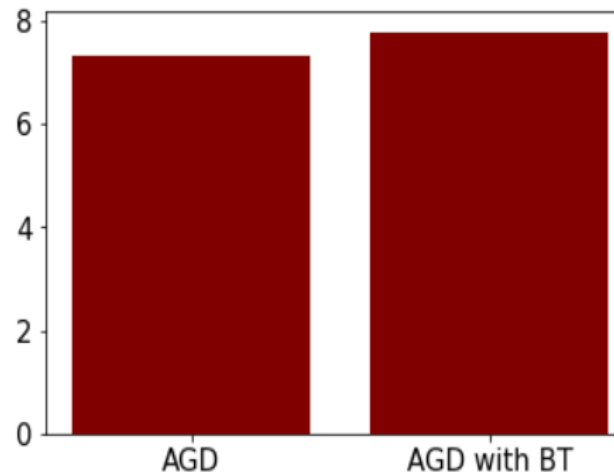
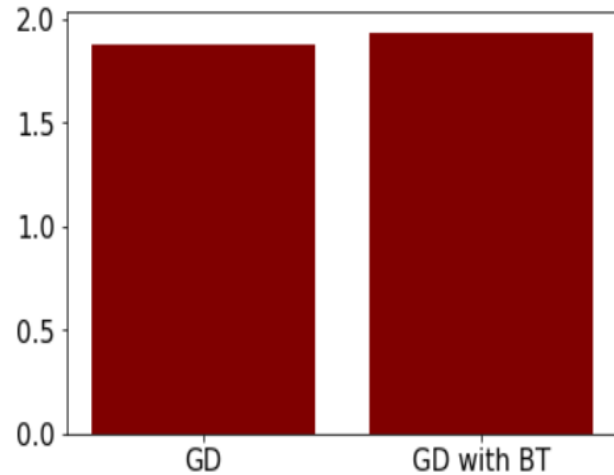
Fitting time of GD with BT: 1.936819076538086

Fitting time of AGD: 7.328400373458862

Fitting time of AGD with BT: 7.77199935913085

Fitting time of Newton: 0.011969327926635742

Fitting time of Newton with BT: 0.003989219665527344

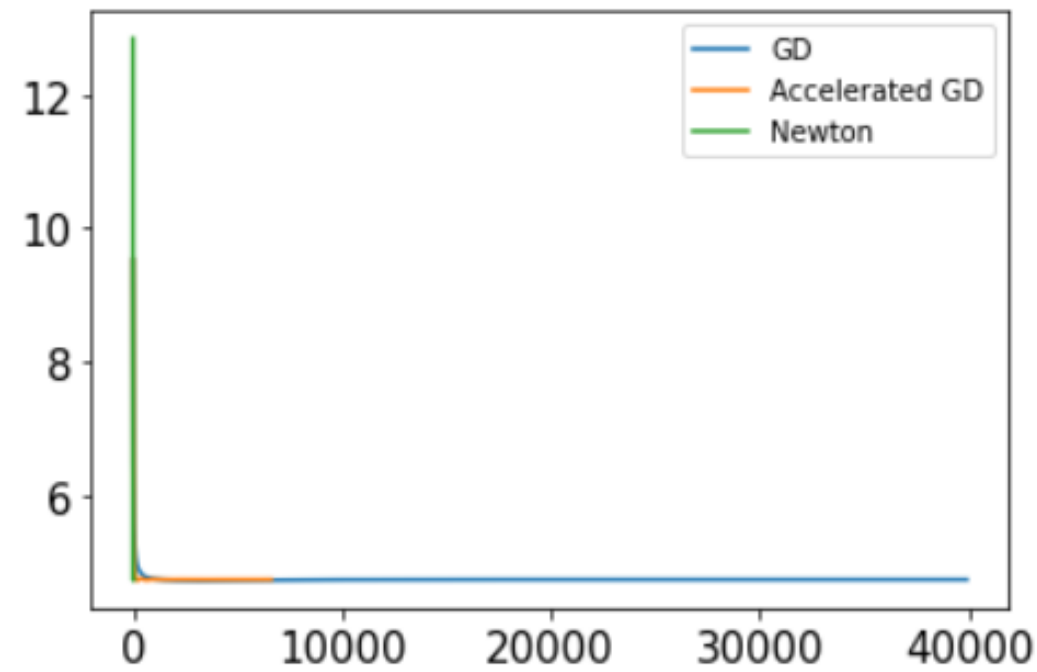


9. So sánh thời gian khi độ dài của grad tiến thời 1 ngưỡng cố định



	R^2 score	Time for fitting (s)	Cost
GD	0.6356	4.04248	4.7441
AGD	0.6350	2.6152	4.75208
Newton	0.6352	0.035954	4.7503

9. So sánh số bước



Total steps with size 0.2: 39888

Total steps with size 0.2: 6606

Total steps of size 1: 9

Fitting time of GD: 10.30033

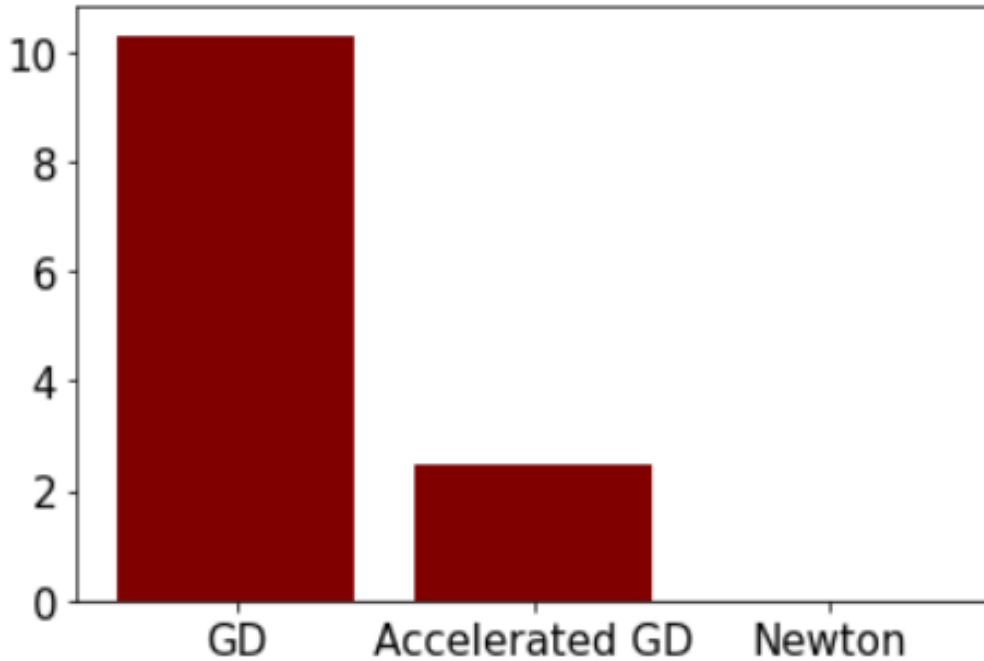
Fitting time of Accelerated GD: 2.49034

Fitting time of Newton: 0.00499

*Với $\text{tol} = 10\text{e-}7$

	Backtracking	Stepsize
GD	False	0.2
AGD	False	0.2
Newton	True	1

9. So sánh thời gian khi độ dài của grad tiến thời 1 ngưỡng cố định



Total steps with size 0.2: 39888

Total steps with size 0.2: 6606

Total steps of size 1: 9

Fitting time of GD: 10.30033

Fitting time of Accelerated GD: 2.49034

Fitting time of Newton: 0.00499

*Với $\text{tol} = 10\text{e-}7$

	Backtracking	Stepsize
GD	False	0.2
AGD	False	0.2
Newton	True	1

Thank you for your time

