Bài 1:

Sử dụng một giá trị cố định cho αk trong quá trình gradient descent có ý nghĩa là tốc độ học được giữ nguyên trong quá trình tối ưu hóa, nếu giá trị cố định này quá nhỏ thì việc dẫn tới hội tụ đến giá trị min là rất lâu, tuy nhiên vẫn có thể tìm được nếu rơi vào trường hợp xấu thì sẽ dẫn tới không thể tìm được giá trị min. Nếu giá trị cố định này quá to thì việc hội tụ sẽ rất nhanh, tuy nhiên nó sẽ không thể chạy đến giá trị min mà chỉ chạy loanh quanh giá trị min mà không thể hội tụ về được.

Ảnh có chứa văn bản, Sơ đồ, biểu đồ, hàng

Mô tả được tạo tự độngẢnh có chứa văn bản, biểu đồ, Sơ đồ, hàng

Mô tả được tạo tự động

Hình nếu giá trị αk quá lớn không thể hội tụ hình αk quá bé hội tụ rất lâu

Đặt αk là một hàm của k trong gradient descent cho phép tốc độ học thích ứng, giúp tăng tốc quá trình hội tụ bằng cách điều chỉnh kích thước bước động động(Chưa hội tụ đến gần giá trị min thì tăng giá trị này lớn để nhanh chóng hội tụ, khi gần tới giá trị hội tụ thì giảm giá trị này xuống để nó hội tụ về đến giá trị min) dựa trên tiến trình của quá trình tối ưu.

Ảnh có chứa văn bản, Sơ đồ, biểu đồ, hàng

Mô tả được tạo tự động

Hình giá trị αk hội tụ theo điều chỉnh thì sẽ thỏa mãn hội tụ

Bài 2:

A, v = φ(x2) - φ(x1) = [0, 2, 2]

B,