



Conjugate Gradient Descent

*E-mail: pima.vn@gmail.com

Mô tả dự án

Trong PiMA lần này, các em sẽ được giới thiệu về phương pháp thông dụng nhất để giải các bài toán tối ưu hàm số với số lượng biến rất lớn: Gradient Descent. **Conjugate Gradient Descent** (CGD) là một phiên bản phức tạp của hơn của Gradient Descent, phỏng theo thuật toán xấp xỉ nghiệm của hệ phương trình tuyến tính. Trong dự án này, chúng ta sẽ tìm hiểu trước CGD cho hàm bậc 2 nhiều biến (quadratic function) và tổng quát hóa ý tưởng này cho một hàm số khả vi bất kì.

Công cụ chính để xây dựng phương pháp này là đại số tuyến tính: sử dụng cơ sở liên hợp (conjugate basis) tương ứng với tích vô hướng (inner product) xác định bởi 1 ma trận đối xứng xác định dương (symmetric positive-definite).

Trong dự án này, các bạn sẽ tìm hiểu nền tảng lý thuyết của CGD và áp dụng công cụ này vào việc tối ưu hóa.

Câu hỏi gợi ý

Hãy tìm hiểu và trình bày chi tiết mô hình CGD, tham khảo các câu hỏi gợi ý như sau.

- (1) Trình bày lại thuật toán gradient descent cơ bản và thuật toán CGD.
- (2) Tại sao tối ưu bậc 2 tương đương giải hệ phương trình tuyến tính? Tại sao lại gọi là conjugate gradient descent?
- (3) Learning rate trong conjugate gradient được chọn như thế nào? Vì sao?
- (4) Tổng quát hóa thuật toán trên cho một hàm số khả vi.
- (5) Phân tích ưu, nhược điểm của CGD.
- (6) Lập trình thuật toán CGD cho hàm Rosenbrock và vẽ đồ thị mô phỏng quá trình cập nhật. So sánh với đồ thị khi sử dụng phương pháp gradient cơ bản.
- (7) Lập trình thuật toán CGD cho một hàm khả vi bất kì khi biết gradient của hàm số đó.

Một số từ khóa: Optimization, Gradient Descent, Inner Product, Quadratic Function, Conjugate Vector, Symmetric Positive-Definite matrix.

Tham Khảo

- [1] Linh Tran, Nghia Nguyen, Trung Can. *Bài giảng Đại số Tuyến tính*, PiMA 2018.
- [2] https://en.wikipedia.org/wiki/Conjugate_gradient_method
- [3] <http://andrew.gibiansky.com/blog/machine-learning/conjugate-gradient/>
- [4] https://en.wikipedia.org/wiki/Rosenbrock_function