

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC KHOA HỌC



TRẦN VĂN PHƯỢNG

VỀ BÀI TOÁN TỐI ƯU
TRONG HỌC ĐỘ TƯƠNG TỰ

LUẬN VĂN THẠC SĨ TOÁN HỌC

THÁI NGUYÊN - 2019

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC KHOA HỌC



TRẦN VĂN PHƯỢNG

VỀ BÀI TOÁN TỐI ƯU
TRONG HỌC ĐỘ TƯƠNG TỰ

Chuyên ngành: Toán ứng dụng

Mã số : 8 46 01 12

LUẬN VĂN THẠC SĨ TOÁN HỌC

NGƯỜI HƯỚNG DẪN KHOA HỌC

TS. Nguyễn Thanh Sơn

THÁI NGUYÊN - 2019

Mục lục

Bảng ký hiệu	1
Mở đầu	2
Chương 1 Bài toán tối ưu trong không gian hữu hạn chiều	6
1.1 Sơ lược về bài toán tối ưu	6
1.1.1 Bài toán tối ưu	6
1.1.2 Khái quát bài toán tối ưu có ràng buộc	8
1.1.3 Tối ưu hàm mục tiêu bậc hai với ràng buộc bất đẳng thức	10
1.2 Một số phương pháp giải bài toán tối ưu	11
1.2.1 Phương pháp Newton	11
1.2.2 Phương pháp giảm sâu nhất	13
1.2.3 Phương pháp hàm chẵn logarith	15
1.2.4 Phương pháp chiếu gradient	18
Chương 2 Bài toán học độ tương tự	21
2.1 Bài toán học độ tương tự và các kiến thức liên quan	21
2.1.1 Một số kiến thức liên quan	21
2.1.2 Bài toán học độ tương tự	25
2.1.3 Tính lồi của bài toán	26
2.1.4 Khoảng cách Mahalanobis	27
2.2 Phương pháp giải bài toán học độ tương tự	28
2.2.1 Trường hợp khoảng cách Euclide có trọng số	28

2.2.2	Phương pháp chiều gradient cho bài toán học độ tương tự	29
2.2.3	Ví dụ số	32
	Kết luận	36
	Tài liệu tham khảo	37

Bảng ký hiệu

H	không gian Hilbert thực
∇f	gradient của hàm số, grad f
$\nabla^2 f$	Hessian của hàm số f và là ma trận cỡ $n \times n$
$\ A\ $	chuẩn Euclid của ma trận A
$\lambda(A)$	các giá trị riêng của A
$A \geq 0$	ma trận nửa xác định dương
$A > 0$	ma trận xác định dương
x^*	điểm cực tiểu hay cực tiểu
$f(x^*)$	giá trị cực tiểu

Mở đầu

Thế giới đang bước vào Cuộc cách mạng khoa học lần thứ tư với AI (Artificial Intelligence - Trí tuệ nhân tạo) và IoT (Internet of Things - Internet vạn vật) đem đến những đột phá bất ngờ về công nghệ. Vai trò của nó đối với một quốc gia, một vùng lãnh thổ lớn đến mức được nhận định, rằng ai dẫn đầu về công nghệ này sẽ chiến thắng trong cuộc cạnh tranh về công nghệ, kinh tế...

Trong thực tế, Trí tuệ nhân tạo mà cụ thể hơn nữa là Học máy (Machine Learning) đã là một ngành con phát triển từ lâu của khoa học máy tính. Luận văn này sẽ xét một bài toán nhỏ trong lĩnh vực rộng lớn này dưới góc độ toán học, đó là *Học độ tương tự*.

Trong một không gian Euclide, hay tổng quát hơn là trong một không gian metric, khái niệm khoảng cách được dùng để đo khoảng cách giữa hai đối tượng. Đối tượng là trùng nhau nếu khoảng cách bằng 0, ở gần nhau nếu khoảng cách nhỏ và xa nhau nếu khoảng cách lớn.

Bây giờ, ta xét một tập hợp tổng quát hơn như tập các hình chụp mặt người, hay tập các văn bản. Giả sử ta có thể biến đổi các đối tượng đó thành các đối tượng toán học như véc tơ hoặc ma trận. Ta cần phải xây dựng một phép đo để có thể phân biệt được hai nhóm đối tượng tương tự nhau và khác nhau. Về mặt định lượng, hai đối tượng tương tự nhau nếu có khoảng cách (theo phép đo vừa được xây dựng) nhỏ và hai đối tượng khác nhau sẽ có khoảng cách lớn.

Câu hỏi tiếp theo là xây dựng phép đo đó như thế nào? Ý tưởng tự

nhiên là khái quát hóa khoảng cách Euclide. Ta có với $x, y \in \mathbb{R}^n$ thì

$$\|x - y\|_E = \sqrt{\|x - y\|^2} = \sqrt{(x - y)^T(x - y)} = \sqrt{(x - y)^T I(x - y)},$$

trong đó I là một ma trận đơn vị. Bây giờ, ta thay I bằng một ma trận đối xứng nửa xác định dương A và định nghĩa

$$\|x - y\|_A = \sqrt{(x - y)^T A(x - y)}. \quad (0.0.1)$$

Lưu ý rằng khi đó (0.0.1) chỉ là một giả khoảng cách, tức là hai điểm khác nhau có thể có khoảng cách bằng 0. Như vậy, việc xây dựng khoảng cách được quy về việc tìm một ma trận đối xứng nửa xác định dương.

Nếu không có gợi ý gì thì đây là một bài toán không có lời giải. Ở góc độ Học máy, muốn máy phân biệt được thế nào là hai đối tượng là tương tự, thế nào là không tương tự thì ta phải dạy nó. Thông tin gợi ý ở đây là việc cho trước hai tập con S và D của không gian các đối tượng (giả sử là \mathbb{R}^n) mà trong đó S chứa những đối tượng tương tự nhau, còn D chứa những đối tượng không tương tự. Một phép đo tốt, trong trường hợp này đặc trưng bởi ma trận A , phải thỏa mãn ba điều:

- i) Khoảng cách giữa các đối tượng thuộc S theo ma trận A càng nhỏ càng tốt;
- ii) Khoảng cách giữa các đối tượng thuộc D theo ma trận A phải tương đối lớn;
- iii) Ma trận A phải thỏa mãn các điều kiện để xây dựng được giả khoảng cách, tức là A phải đối xứng và nửa xác định dương.

Những gợi ý trên đã dẫn đến bài toán tối ưu sau:

$$\arg \min_A \sum_{(x_i, x_j) \in S} \|x_i - x_j\|_A^2 \quad (0.0.2)$$

sao cho

$$\sum_{(x_i, x_j) \in D} \|x_i - x_j\|_A \geq 1, \quad A \geq 0. \quad (0.0.3)$$

Mục đích của đề tài là trình bày việc giải bài toán tối ưu nảy sinh trong học độ tương tự (0.0.2), (0.0.3).

Nội dung của đề tài luận văn được trình bày trong hai chương.

Chương 1. Bài toán tối ưu trong không gian hữu hạn chiều

Nội dung chính của chương này là các kiến thức về tối ưu hóa. Chúng tôi nhắc lại một cách sơ lược các khái niệm cơ bản của bài toán tối ưu và một số phương pháp cho bài toán tối ưu không ràng buộc. Trong đó, chúng tôi đi sâu vào trình bày chi tiết phương pháp chiếu gradient sẽ dùng ở Chương 2. Tài liệu tham khảo chính cho chương này là [2], [4].

Chương 2. Bài toán học độ tương tự

Trong chương này, luận văn trình bày những chủ đề khái quát về bài toán học độ tương tự. Sau đó, chúng tôi đi vào trình bày chi tiết bài toán Học độ tương tự theo loạt. Còn một số chủ đề rất thú vị khác như Học độ tương tự online, Học độ tương tự dựa trên lý thuyết thông tin đã không thể được trình bày do khuôn khổ có hạn của luận văn cũng như sự hạn chế về thời gian và năng lực. Sau cùng, chương này trình bày phương pháp chiếu gradient cho Bài toán học độ tương tự và ví dụ số minh họa cho phương pháp đó. Chương này tham khảo các tài liệu [3], [5].

Luận văn được hoàn thành tại Trường Đại học Khoa học - Đại học Thái Nguyên. Trong quá trình học tập và thực hiện luận văn này, Trường Đại học Khoa học đã tạo mọi điều kiện tốt nhất để tác giả học tập, nghiên cứu. Tác giả xin được bày tỏ lòng biết ơn chân thành đến các thầy, cô trong khoa Toán - Tin, trong Trường Đại học Khoa học - Đại học Thái Nguyên. Đặc biệt, tác giả xin bày tỏ lòng biết ơn sâu sắc tới TS. Nguyễn Thanh Sơn - Người đã tận tình hướng dẫn tác giả hoàn

thành luận văn này.

Tác giả cũng xin được gửi lời cảm ơn tới Ban giám hiệu trường THPT Nguyễn Đăng Đạo và tập thể các thầy cô giáo trong tổ Toán-Tin của Trường đã tạo điều kiện giúp đỡ tác giả trong thời gian tác giả tham gia học cao học.

Thái Nguyên, tháng 04 năm 2019

Tác giả luận văn

Trần Văn Phương

Chương 1

Bài toán tối ưu trong không gian hữu hạn chiều

Nội dung chính của chương này là các kiến thức về tối ưu hóa. Chúng tôi nhắc lại một cách sơ lược các khái niệm cơ bản của bài toán tối ưu và một số phương pháp cho bài toán tối ưu bao gồm bài toán không ràng buộc và bài toán có ràng buộc. Trong đó, chúng tôi đi sâu vào trình bày chi tiết phương pháp chiếu gradient sẽ dùng ở Chương 2. Tài liệu tham khảo chính cho chương này là [2], [4].

1.1 Sơ lược về bài toán tối ưu

Mục này sẽ trình bày khái niệm, kết quả cơ bản để có cái nhìn khái quát về bài toán tối ưu.

1.1.1 Bài toán tối ưu

Cho $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Tìm cực tiểu địa phương x^* của f , nghĩa là,

$$f(x^*) \leq f(x), \forall x \in \mathcal{U}_{x^*}, \quad (1.1.1)$$

trong đó, \mathcal{U}_{x^*} là lân cận địa phương nào đó của x^* . Để ngắn gọn, ta viết

$$\min_x f(x). \quad (1.1.2)$$