ĐẠI HỌC HUẾ
TRƯỜNG ĐẠI HỌC SƯ PHẠM

PHAN THỊ LIỄU

ỨNG DỤNG QUY HOẠCH LỒI VÀO KĨ THUẬT MÁY VÉC-TƠ TỰA

LUẬN VĂN THẠC SĨ KHOA HỌC

CHUYÊN NGÀNH: GIẢI TÍCH TOÁN HỌC

MÃ SỐ: 8460102

Người hướng dẫn: PGS.TS HUỲNH THẾ PHÙNG

Huế - 2019

LỜI CAM ĐOAN

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu và kết quả ghi trong luận văn là trung thực, được các đồng tác giả cho phép sử dụng và chưa từng được công bố trong bất kì một công trình nào khác.

PHAN THỊ LIỄU

LỜI CẢM ƠN

Luận văn này được hoàn thành dưới sự hướng dẫn của Thầy giáo, PGS.TS Huỳnh Thế Phùng. Tôi xin bày tỏ lòng biết ơn sâu sắc và sự kính trọng đối với Thầy. Thầy đã tận tình hướng dẫn, giúp đỡ tôi trong quá trình học tập cũng như hoàn thành luận văn này.

Tôi xin gửi lời cảm ơn đến quý Thầy cô khoa Toán, các Thầy ở Đại Học Huế đã dạy dỗ và truyền đạt kiến thức cho tôi trong suốt quá trình học tập.

Tôi xin chân thành cảm ơn Ban giám hiệu trường Đại Học Sư Phạm Huế, phòng Đào tạo sau Đại học, khoa Toán trường Đại Học Sư Phạm Huế đã tạo điều kiện cho tôi trong suốt quá trình học.

Cuối cùng, tôi xin cảm ơn gia đình, bạn bè, các anh chị Cao học Toán khóa XXVI trường Đại Học Sư Phạm Huế vì sự động viên , giúp đỡ trong quá trình học tập vừa qua.

Ngày 5 tháng 10 năm 2019

Học viên thực hiện

PHAN THỊ LIỄU

Mục lục

Trang phụ bìa **i**

Lời cam đoan **ii**

Lời cảm ơn **iii**

Mục lục **1**

Lời mở đầu **3**

1 Bài toán tối ưu lồi **4**

1.1 Điều kiện tối ưu đối với bài toán tối ưu tổng quát 4

1.2 Điều kiện cực trị 8

1.3 Bài toán qui hoạch toàn phương 9

1.4 Bài toán đối ngẫu của bài toán qui hoạch toàn phương lồi 13

2 Máy véc-tơ tựa (SVM) **16**

2.1 Phát biểu bài toán SVM cơ bản (hai lớp) 16

2.2 Xây dựng bài toán tối ưu cho SVM 19

2.3 Bài toán đối ngẫu của SVM 23

2.3.1 Kiểm tra tiêu chuẩn Slater 24

2.3.2 Hàm Lagrange của bài toán SVM 24

2.3.3 Hàm đối ngẫu Lagrange 24

2.3.4 Bài toán đối ngẫu Lagrange 25

2.3.5 Điều kiện KKT 26

2.4 Bài toán lề mềm 27

2.4.1 Phân tích toán học 29
2.4.2 Bài toán đối ngẫu Lagrange 31
2.4.3 Bài toán tối ưu không ràng buộc cho SVM lề mềm 34

3 Máy véc-tơ tựa nhiều lớp 39
3.1 Xây dựng hàm mất mát cho SVM nhiều lớp 40
3.1.1 Chính quy hóa 43
3.1.2 Chọn giá trị $\triangle$ 43
3.2 Tính toán hàm mất mát và đạo hàm 44

Kết luận 47

Tài liệu tham khảo 49

LỜI MỞ ĐẦU

Phân lớp dữ liệu hay phân cụm dữ liệu là một bài toán xuất hiện rất phổ biến trong đời sống thực tế, trở thành một lĩnh vực nghiên cứu quan trọng trong công nghệ thông tin mà được phát triển hoàn toàn dựa trên nền tảng là lý thuyết tối ưu lồi. Việc tìm hiểu bài toán này rõ ràng là rất hấp dẫn và là cơ hội để tôi có thể vận dụng các kiến thức đã học từ giải tích lồi và lý thuyết các bài toán cực trị trong khảo sát các bài toán thực tế. Vì vậy được sự đồng ý hướng dẫn của PGS.TS Huỳnh Thế Phùng chúng tôi chọn đề tài "Ứng dụng quy hoạch lồi vào kĩ thuật máy véc-tơ tựa" cho luận văn thạc sĩ của mình.

Kĩ thuật máy véc-tơ tựa (SVM) là một kĩ thuật phân lớp do Vapnik và Chervonenkis xây dựng năm 1995 và có nhiều tiềm năng phát triển về mặt lý thuyết cũng như ứng dụng trong thực tế. Nó là công cụ đắc lực cho các bài toán về xử lý ảnh, phân loại văn bản, nhận dạng giọng nói, chữ viết tay, nhận dạng bệnh,...

Với vấn đề đặt ra luận văn bao gồm các nội dung sau:

Chương 1: Bài toán tối ưu lồi. Chương này nhắc lại một số kết quả cơ bản về bài toán tối ưu lồi, bài toán quy hoạch toàn phương và bài toán đối ngẫu của nó.

Chương 2: Máy véc-tơ tựa (SVM). Chương này phát biểu bài toán SVM cơ bản (hai lớp), xây dựng bài toán tối ưu cho SVM, bài toán đối ngẫu và bài toán lề mềm của SVM.

Chương 3: Máy véc-tơ tựa nhiều lớp. Chương này xây dựng hàm mất mát cho SVM nhiều lớp; tính toán hàm mất mát và đạo hàm.

Do hạn chế về mặt thời gian và kiến thức nên luận văn không tránh khỏi những sai sót, rất mong nhận được sự đóng góp ý kiến của quý thầy cô và các bạn để luận văn được hoàn thiện hơn.

Huế, tháng 10 năm 2019

PHAN THỊ LIỄU

Chương 1

Bài toán tối ưu lồi

Chương này nhằm nhắc lại một số kết quả cơ bản về bài toán tối ưu lồi mà nó được sử dụng trong luận văn. Kết quả được khảo sát chủ yếu bởi bài toán quy hoạch toàn phương lồi.

1.1 Điều kiện tối ưu đối với bài toán tối ưu tổng quát

Bài toán tối ưu tổng quát có dạng như sau:

$$\mathcal{P}(\mathbb{R}^n; g; h; f) : \begin{cases} f(x) \to \min \\ g(x) \leq 0, \\ h(x) = 0, \end{cases} \tag{1.1}$$

trong đó, $g : \mathbb{R}^n \to \mathbb{R}^m$, $h : \mathbb{R}^n \to \mathbb{R}^k$, $f : \mathbb{R}^n \to \mathbb{R}$.

Tập hợp sau đây

$$M := \{x \in \mathbb{R}^n | \; g(x) \leq 0, \; h(x) = 0\}, \tag{1.2}$$

được gọi là tập chấp nhận được của bài toán.

Các ràng buộc $x \in \mathbb{R}^n$, $g(x) \leq 0$, $h(x) = 0$ lần lượt gọi là ràng buộc tập, ràng buộc bất đẳng thức và ràng buộc đẳng thức của bài toán. Rõ ràng nếu f, g_i, h_j là các hàm affine thì $\mathcal{P}(\mathbb{R}^n; g, h, f)$ là bài toán qui hoạch tuyến tính; nếu f, g_i là các hàm lồi, h_j là hàm affine thì đó là bài toán qui hoạch lồi; nếu f, g_i, h_j khả vi liên tục thì ta có bài toán qui hoạch trơn.

Bổ đề 1.1.1. *Cho $a \in \mathbb{R}^n$, $b \in \mathbb{R}^k$. Lúc đó,*

$$\begin{aligned}
[a^T x \le 0, \forall x \in \mathbb{R}^m_+] &\Leftrightarrow [a \le 0].\\
[b^T y \le 0, \forall y \in \mathbb{R}^k] &\Leftrightarrow [b = 0].\\
[a^T x + b^T y \le 0, \forall (x, y) \in \mathbb{R}^m_+ \times \mathbb{R}^k] &\Leftrightarrow [(a \le 0) \text{ và } (b = 0)].\\
[a^T x \le \alpha, \forall x \in \mathbb{R}^m] &\Leftrightarrow [(a = 0) \text{ và } (\alpha \ge 0)].
\end{aligned}$$

Bổ đề 1.1.2. *Cho $a = (a_1, a_2, ..., a_m) \in \mathbb{R}^m$, $b = (b_1, b_2, ..., b_k) \in \mathbb{R}^k$. Lúc đó:*

$$\begin{aligned}
\sup_{x \in \mathbb{R}^m_+} a^T x &= \begin{cases} 0, & \text{nếu } a \le 0,\\ +\infty, & \text{nếu tồn tại } a_i > 0. \end{cases}\\
\sup_{y \in \mathbb{R}^k} b^T y &= \begin{cases} 0, & \text{nếu } \quad b = 0,\\ +\infty, & \text{nếu } \quad b \ne 0. \end{cases}\\
\sup [a^T x + b^T y] &= \begin{cases} 0, & \text{nếu } \quad (a \le 0) \text{ và } (b = 0),\\ +\infty, & \text{nếu } \quad (b \ne 0) \ (\text{hoặc tồn tại } a_i > 0). \end{cases}
\end{aligned}$$

Từ bổ đề trên, với tập M được cho bởi (1.2),

$$x \in M \Leftrightarrow \sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} (\lambda^T g(x) + \mu^T h(x)) = 0. \tag{1.3}$$

Do đó, với mỗi $x \in \mathbb{R}^m$, ta có

$$\sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \{f(x) + \lambda^T g(x) + \mu^T h(x)\} = \begin{cases} f(x), & \text{nếu} \quad x \in M;\\ +\infty, & \text{nếu} \quad x \ne M. \end{cases}$$

Như vậy, nếu đặt

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x) \qquad x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^k. \tag{1.4}$$

thì

$$\tilde{f}(x) := \sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \mathcal{L}(x, \lambda, \mu) = \begin{cases} f(x), & \text{nếu} \quad x \in M,\\ +\infty, & \text{nếu} \quad x \in \mathbb{R}^n \backslash M. \end{cases} \tag{1.5}$$

Từ đó, bài toán $\mathcal{P}(M; f)$ đưa được về bài toán tương đương dạng minimax theo cách sau đây:

$$\inf_{x \in M} f(x) = \inf_{x \in \mathbb{R}^n} \tilde{f}(x) = \inf_{x \in \mathbb{R}^n} \sup_{(\lambda,\mu) \in \mathbb{R}^m_+ \times \mathbb{R}^k} \mathcal{L}(x, \lambda, \mu). \tag{1.6}$$