MÁY VÉC-TƠ HỖ TRỢ

HAI LỚP, KHẢ TÁCH TUYẾN TÍNH

· · · BÀI TOÁN ĐỐI NGẪU · · ·

Lê Thành Sách

✉ ltsach@hcmut.edu.vn

Khoa Khoa học & Kỹ thuật Máy tính

Trường Đại học Bách Khoa - ĐHQG Tp.HCM

Tp.HCM. Ngày 21 tháng 9 năm 2019

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Phương pháp mở rộng

Tổng kết

Mục lục

1. Bài toán gốc: ôn lại
2. Kiến thức liên quan
3. Bài toán đối ngẫu
4. Phương pháp mở rộng
5. Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán gốc: ôn lại

Mục lục

2 Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán gốc

Bài toán tối ưu gốc: $\mathcal{P}_{primal}$

$$
\begin{aligned}
\boldsymbol{w}^*, b^* &= \underset{\boldsymbol{w},b}{\operatorname{argmin}} \frac{1}{2}||\boldsymbol{w}||^2 \\
\text{s.t:} \quad & 1 - t_n(\boldsymbol{w}^T\boldsymbol{x}_n + b) \leq 0 \\
& \forall n = 1, 2, \ldots, N
\end{aligned} \tag{1.1}
$$

Dạng cơ bản $\mathcal{P}_{basic}$

$$
\begin{aligned}
p^* \quad &= \underset{\boldsymbol{x}}{\text{minimize}}\, f_0(\boldsymbol{x}) \\
\text{s.t:} \quad & f_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \ldots, m \\
& h_i(\boldsymbol{x}) = 0, \qquad i = 1, \ldots, p
\end{aligned} \tag{1.2}
$$

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán gốc

Mục lục
4 Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Phương pháp mở rộng
Tổng kết

Bài toán tối ưu gốc: $\mathcal{P}_{primal}$

$$\begin{aligned} \boldsymbol{w}^*, b^* &= \underset{\boldsymbol{w}, b}{\operatorname{argmin}} \frac{1}{2}||\boldsymbol{w}||^2 \\ \text{s.t:} \quad & 1 - t_n(\boldsymbol{w}^T\boldsymbol{x}_n + b) \leq 0 \\ & \forall n = 1, 2, \ldots, N \end{aligned}$$

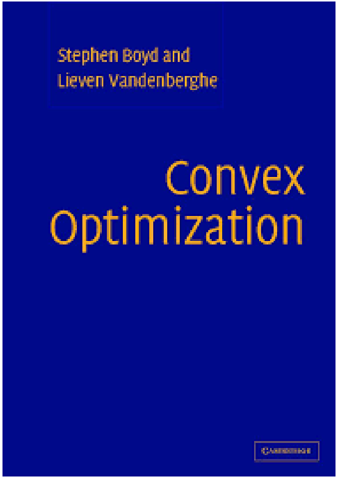
- Hàm mục tiêu: $f_0(\boldsymbol{w}, b) = \frac{1}{2}||\boldsymbol{w}||^2$
- Các ràng buộc dạng bất đẳng thức: $f_n(\boldsymbol{w}, b) = 1 - t_n(\boldsymbol{w}^T\boldsymbol{x}_n + b)$[1]
- Các ràng buộc dạng đẳng thức: **không có**
- Hàm mục tiêu và các hàm ràng buộc: **lồi và khả vi**

[1] $f_n(\boldsymbol{w}, b) \leq 0, \forall n = 1, 2, \ldots, N$

Kiến thức liên quan

Mục lục
Bài toán gốc: ôn lại
5 Kiến thức liên quan
Bài toán đối ngẫu
Phương pháp mở rộng
Tổng kết

Tối ưu lồi



Hình 2.1: Convex Optimization của Stephen Boyd và Lieven Vandenberghe. Cần đọc Chương 5.

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

6 Kiến thức liên quan

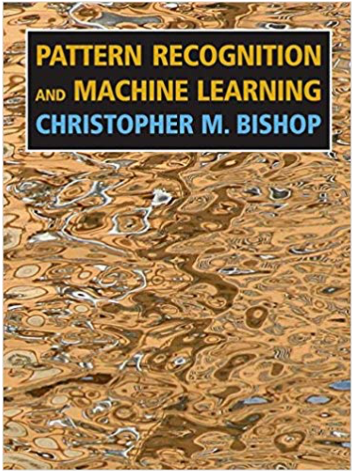
Bài toán đối ngẫu

Phương pháp mở rộng

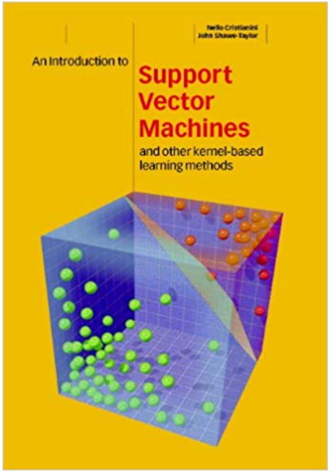
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Máy véctơ hỗ trợ



Hình 2.2: (a) **Machine Learning**, Bishop, (Phần 7.1)



Hình 2.3: (b) **Support Vector Machines**, Nello et al., (Chương 6)

Bài toán đối ngẫu

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
8 Bài toán đối ngẫu
Nguyên tắc
Lagrangian
Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Nguyên tắc

1. Thành lập hàm **Lagrangian** $\mathcal{L}$ để liên kết các biến gốc với các biến đối ngẫu.
   1. Mỗi hàm ràng buộc ở bài toán gốc được nhân với một hệ số, được gọi là biến đối ngẫu
   2. $\mathcal{L}$ là tổng của hàm mục tiêu gốc với **tổ hợp tuyến tính** gồm các hàm ràng buộc gốc với hệ số là các **biến đối ngẫu**
2. Dựa vào tính chất của hàm $\mathcal{L}$ và của bài toán gốc, chuyển sang bài toán đối ngẫu bằng cách:
   1. Xây dựng **hàm đối ngẫu** (hàm thu được khi cực tiểu hàm $\mathcal{L}$ theo các biến gốc)
   2. Bài toán đối ngẫu: **cực đại hóa hàm đối ngẫu**
3. Sử dụng tính chất của hệ điều kiện **KKT** và **Slater** suy ra nghiệm của bài toán gốc.

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

9 Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu
Lagrangian

$$\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}||\mathbf{w}||^2 - \sum_{n=1}^{N} \alpha_n \{t_n(\mathbf{w}^T \mathbf{x}_n + b) - 1\} \quad (3.1)$$

Tên gọi

- $\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha})$: Hàm Lagrangian[1]
- $\alpha_n; n = 1, 2, \cdots, N$: Các nhân tử Lagrange[2]
- Lagrangian liên kết các biến của bài toán gốc và các biến mới vào một hàm

[1] còn gọi: **Lagrangian**
[2] còn gọi: biến đối ngẫu (hay **dual variables**)

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Lagrangian

$$\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}||\mathbf{w}||^2 - \sum_{n=1}^{N} \alpha_n \{t_n(\mathbf{w}^T \mathbf{x}_n + b) - 1\}$$

Tính chất

- Lagrangian là **convex** theo $\mathbf{w}, b$: hàm bậc 2 theo $\mathbf{w}, b$ (vì ma trận cho p.tử bậc hai $||\mathbf{w}||^2$ là bán định dương)
- Lagrangian là **concave** theo $\boldsymbol{\alpha}$; là hàm **affine** theo $\boldsymbol{\alpha}$ (vì $\mathcal{L}$ là tổ hợp tuyến tính theo α_n)

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

11 Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn

Lê Thành Sách

Bài toán đối ngẫu

Lagrangian

$$\mathcal{L}(\mathbf{w}, b, \alpha) = \underbrace{\frac{1}{2}||\mathbf{w}||^2 - \underbrace{\sum_{n=1}^{N} \underbrace{\alpha_n}_{\text{ép buộc:}\geq 0} \underbrace{\{t_n(\mathbf{w}^T\mathbf{x}_n + b) - 1\}}_{\text{ràng buộc gốc:}\geq 0}}_{\geq 0}}_{\leq \frac{1}{2}||\mathbf{w}||^2}$$

Tính chất; với ràng buộc $\alpha_n \geq 0$; $\mathbf{w}^*, b^*$: là điểm tối ưu của bài toán gốc

- $\mathcal{L}(\mathbf{w}^*, b^*, \alpha)$: là chặn dưới của hàm mục tiêu gốc tại điểm tối ưu: $\mathcal{L}(\mathbf{w}^*, b^*, \alpha) \leq \frac{1}{2}||\mathbf{w}^*||^2$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

12 Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Lagrangian: như "cost function"

$$\underbrace{\underbrace{\mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}||\mathbf{w}||^2 + \sum_{n=1}^{N} \underbrace{\alpha_n}_{\text{ép buộc:}\geq 0} \underbrace{\{1 - t_n(\mathbf{w}^T \mathbf{x}_n + b)\}}_{\text{đúng}:\leq 0;\quad \text{sai}:>0}}_{E_{data}:\text{cost, chí phí}}}_{\text{Tổng chi phí} = \frac{1}{2}||\mathbf{w}||^2 + E_{data}}$$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Lagrangian: như "cost function"

- $\mathcal{L}$: như hàm chi phí; **vừa tối thiểu** $||\boldsymbol{w}||^2$ **và** b**, vừa phải giảm chi phí tính trên dữ liệu huấn luyện** E_{data}
- $E_{data} \leq 0$: khi phân loại đúng $\Rightarrow$ càng âm càng tốt, càng dương càng xấu
- $\alpha_n \geq 0 \, (n = 1, 2, \cdots, N) \Rightarrow$ Cần cực đại hóa theo các α_n

Cách thực hiện

- Tối thiểu $\mathcal{L}$ theo $\boldsymbol{w}$ và b trước $\Rightarrow$ Hàm đối ngẫu.
- Cần cực đại hàm kết quả theo $\boldsymbol{\alpha} \Rightarrow$ Nghiệm của bài toán đối ngẫu $(\boldsymbol{\alpha}^*)$

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán

Bài toán gốc là: Tối ưu lồi (convex optimization) , vì:

1. Hàm mục tiêu gốc:
   $f_0(\mathbf{w}, b) = \frac{1}{2}||\mathbf{w}||^2$: **lồi**
2. Các hàm ràng buộc dạng bất đẳng thức:
   $f_n(\mathbf{w}, b) = 1 - t_n\{\mathbf{w}^T\mathbf{x}_n + b\}; n = 1, 2, \ldots, N$[1]: **lồi**
3. Các hàm ràng buộc dạng đẳng thức: có dạng **affine**; bài toán gốc không chứa dạng ràng buộc này.

[1] $f_n(\mathbf{w}, b) \leq 0$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán

Bài toán tối ưu gốc có hai tính chất:

1. **tối ưu lồi**
2. các hàm $f_0(\mathbf{w}, b)$ và $f_n(\mathbf{w}, b)$ là **khả vi**

$\Rightarrow$ Thỏa mãn **hệ điều kiện KKT**[1] sau đây

[1] Karush–Kuhn–Tucker: Kết qủa từ Karush (1939) và Kuhn–Tucker (1951)

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán

Hệ điều kiện KKT

(KKT-1) Điều kiện dừng:	$\nabla_{\mathbf{w},b}\mathcal{L}(\tilde{\mathbf{w}}, \tilde{b}, \tilde{\alpha}) = 0$
(KKT-2) Ràng buộc gốc:	$1 - t_n\{\tilde{\mathbf{w}}^T\mathbf{x}_n + \tilde{b}\} \leq 0;$ $n = 1, 2, \ldots, N$
(KKT-3) Ràng buộc biến đối ngẫu:	$\tilde{\alpha}_n \geq 0;$ $n = 1, 2, \ldots, N$
(KKT-4) Điều kiện bù:	$\tilde{\alpha}_n\{1 - t_n(\tilde{\mathbf{w}}^T\mathbf{x}_n + \tilde{b})\} = 0;$ $n = 1, 2, \ldots, N$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu
- Nguyên tắc
- Lagrangian
- 17 Xây dựng bài toán
- Hàm đối ngẫu
- Cực đại hàm đối ngẫu
- Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán

Ý NGHĨA CỦA HỆ ĐIỀU KIỆN KKT

NẾU: $(\tilde{w}, \tilde{b}, \tilde{\alpha})$ thoải mãn KKT:

THÌ:

1. $(\tilde{w}, \tilde{b})$ là nghiệm của bài toán gốc; nghĩa là, $(\tilde{w}, \tilde{b}) \equiv (w^*, b^*)$
2. $(\tilde{\alpha})$ là nghiệm của bài toán đối ngẫu[1] ; nghĩa là, $(\tilde{\alpha}) \equiv (\alpha^*)$
3. **duality gap** $= 0$; nghĩa là, $\mathcal{L}(w^*, b^*, \alpha^*) = \frac{1}{2}||w^*||^2$

[1] trình bày sau

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

18 Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán: nguyên tắc

Từ **Điều kiện dừng** :

1. $\Rightarrow$ Tìm hàm $g(\alpha) = \mathcal{L}(\tilde{w}, \tilde{b}, \alpha)$; với $\tilde{w}, \tilde{b}$ là các điểm dừng[1]; hàm $g(\alpha)$ được gọi làm **hàm đối ngẫu**

Từ **duality gap** $= 0$:

2. $\Rightarrow$ **maximize** hàm $g(\alpha)$ để tìm α^*. Bài toán tối ưu này được gọi là **bài toán đối ngẫu**[2]

[1] là điểm tại đó đạo hàm bằng 0

[2] vì từ nghiệm bài toán này có thể suy ra nghiệm bài toán gốc

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

19 Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu
Xây dựng bài toán: tìm hàm đối ngẫu

- Tính đạo hàm:

$$\frac{\partial \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})}{\partial \boldsymbol{w}} = \boldsymbol{w} - \sum_{n=1}^{N} \alpha_n t_n \boldsymbol{x}_n$$

$$\frac{\partial \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})}{\partial b} = \sum_{n=1}^{N} \alpha_n t_n$$

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
20 Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán: tìm hàm đối ngẫu

- Giải tìm các ràng buộc:

$$\frac{\partial \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})}{\partial \boldsymbol{w}} = \mathbf{0}$$

Công thức của $\boldsymbol{w}$

$$\boldsymbol{w} = \sum_{n=1}^{N} \alpha_n t_n \boldsymbol{x}_n \tag{3.2}$$

Lưu ý: $\boldsymbol{w}$ là tổ hợp tuyến tính của các điểm dữ liệu đầu vào; **nếu tìm được $\boldsymbol{\alpha}$** thì tính $\boldsymbol{w}$ công thức trên.

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
21 Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

Bài toán đối ngẫu

Xây dựng bài toán: tìm hàm đối ngẫu

- Giải tìm các ràng buộc:

$$\frac{\partial \mathcal{L}(\mathbf{w}, b, \boldsymbol{\alpha})}{\partial b} = 0$$

$$\Rightarrow \sum_{n=1}^{N} \alpha_n t_n = 0$$

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán: tìm hàm đối ngẫu

- Hàm đối ngẫu:

$$g(\boldsymbol{\alpha}) = \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})$$

$$\text{với,} \begin{cases} \boldsymbol{w} = \sum_{n=1}^{N} \alpha_n t_n \boldsymbol{x}_n, \ \text{và} \\ \sum_{n=1}^{N} \alpha_n t_n = 0 \end{cases}$$

Mục tiêu

Tìm $\boldsymbol{\alpha}^* = \text{argmax}_{\boldsymbol{\alpha}}\, g(\boldsymbol{\alpha})$

Hoặc $\boldsymbol{\alpha}^* = \text{argmin}_{\boldsymbol{\alpha}} -g(\boldsymbol{\alpha})$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

23 Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Xây dựng bài toán: tìm hàm đối ngẫu

- Hàm đối ngẫu:

$$
\begin{aligned}
g(\boldsymbol{\alpha}) &= \frac{1}{2}\sum_{r=1}^{N}\sum_{c=1}^{N}\alpha_c\alpha_r t_c t_r \mathbf{x}_c^T\mathbf{x}_r - \sum_{r=1}^{N}\alpha_r t_r\left(\sum_{c=1}^{N}\alpha_c t_c \mathbf{x}_c^T\right)\mathbf{x}_r \\
&\quad - b\sum_{n=1}^{N}\alpha_n t_n + \sum_{n=1}^{N}\alpha_n \\
&= -\frac{1}{2}\sum_{r=1}^{N}\sum_{c=1}^{N}\alpha_c\alpha_r t_c t_r \mathbf{x}_c^T\mathbf{x}_r + \sum_{n=1}^{N}\alpha_n \qquad (3.3)
\end{aligned}
$$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

24 Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Hàm đối ngẫu và tính chất

$$g(\boldsymbol{\alpha}) = -\frac{1}{2}\sum_{r=1}^{N}\sum_{c=1}^{N}\alpha_c\alpha_r t_c t_r \underbrace{\mathbf{x}_c^T\mathbf{x}_r}_{\text{tích vô hướng}} + \sum_{n=1}^{N}\alpha_n \tag{3.4}$$

Tính chất:

- $g(\boldsymbol{\alpha})$: **concave**, bất kể bài toán gốc là **convex** hay không
- $-g(\boldsymbol{\alpha})$: **convex**
- Khi $\boldsymbol{\alpha} \geq 0$: $g(\boldsymbol{\alpha}) \leq p^*$ ($g(\boldsymbol{\alpha})$ là cận dưới của hàm mục tiêu gốc)

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

25 Hàm đối ngẫu

Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu (dual problem): hàm mục tiêu

$$
\begin{aligned}
\boldsymbol{\alpha}^* \quad &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2} \sum_{r=1}^{N} \sum_{c=1}^{N} \alpha_c \alpha_r t_c t_r \mathbf{x}_c^T \mathbf{x}_r - \sum_{n=1}^{N} \alpha_n \\
\text{s.t:} \quad & \alpha_n \geq 0; n = 1, 2, \cdots, N \qquad (3.5) \\
& \sum_{n=1}^{N} \alpha_n t_n = 0
\end{aligned}
$$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

26 Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Cực đại hàm đối ngẫu

Bài toán đối ngẫu (dual problem): hàm mục tiêu cho thư viện CVXOPT

$$
\begin{aligned}
\boldsymbol{\alpha}^* &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2}\sum_{r=1}^{N}\sum_{c=1}^{N}\alpha_c\alpha_r t_c t_r \mathbf{x}_c^T\mathbf{x}_r - \sum_{n=1}^{N}\alpha_n \\
\text{s.t:} \quad & -\alpha_n \leq 0; n = 1, 2, \cdots, N \\
& \sum_{n=1}^{N}\alpha_n t_n = 0
\end{aligned}
\tag{3.6}
$$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

27 Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Cực đại hàm đối ngẫu

Bài toán đối ngẫu (dual problem): hàm mục tiêu cho thư viện CVXOPT

$$
\begin{aligned}
\boldsymbol{\alpha}^* \quad &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{K}\boldsymbol{\alpha} + \boldsymbol{p}^T\boldsymbol{\alpha} \\
\text{s.t:} \quad & \boldsymbol{G}\boldsymbol{\alpha} \leq \boldsymbol{h} \\
& \boldsymbol{A}\boldsymbol{\alpha} = b
\end{aligned} \tag{3.7}
$$

Lưu ý: ở ràng buộc, so sánh từng phần tử tương ứng của hai véctơ $\boldsymbol{G}\boldsymbol{\alpha}$ và $\boldsymbol{h}$ với nhau $\Rightarrow$ có N ràng buộc; với các ma trận và véctơ nêu sau đây.

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Cực đại hàm đối ngẫu

$$K_{Gram} = \begin{bmatrix} \mathbf{x}_1^T\mathbf{x}_1 & \mathbf{x}_1^T\mathbf{x}_2 & \cdots & \mathbf{x}_1^T\mathbf{x}_N \\ \mathbf{x}_2^T\mathbf{x}_1 & \mathbf{x}_2^T\mathbf{x}_2 & \cdots & \mathbf{x}_2^T\mathbf{x}_N \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_N^T\mathbf{x}_1 & \mathbf{x}_N^T\mathbf{x}_2 & \cdots & \mathbf{x}_N^T\mathbf{x}_N \end{bmatrix} = \mathbf{X}\mathbf{X}^T \tag{3.8}$$

$$\mathbf{Y} = \begin{bmatrix} y_1y_1 & y_1y_2 & \cdots & y_1y_N \\ y_2y_1 & y_2y_2 & \cdots & y_2y_N \\ \vdots & \vdots & \ddots & \vdots \\ y_Ny_1 & y_Ny_2 \cdots & y_Ny_N & \end{bmatrix} = \mathbf{y}\mathbf{y}^T \text{(outer product)} \tag{3.9}$$

$$\mathbf{K} = \mathbf{K}_{Gram} \odot \mathbf{Y} \tag{3.10}$$

(Element-wise product)

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

29 Cực đại hàm đối ngẫu

Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Cực đại hàm đối ngẫu

$$\boldsymbol{p} = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} \qquad \boldsymbol{x} = \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}$$

- $\boldsymbol{p}$: $N \times 1$; chỉ chứa số -1
- $\boldsymbol{x}$: $N \times 1$; cũng là véctơ $\boldsymbol{\alpha}$

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán tối ưu

Bài toán gốc

$$\boldsymbol{G} = \begin{pmatrix} -1 & & & \\ & -1 & & \\ & & \ddots & \\ & & & -1 \end{pmatrix} \qquad \boldsymbol{h} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- $\boldsymbol{G}$: $N \times N$; các phần tử không xuất hiện là 0; các giá trị theo đường chéo là -1
- $\boldsymbol{h}$: $N \times 1$; gồm các giá trị 0

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Cực đại hàm đối ngẫu

$$
\boldsymbol{A} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_N \end{bmatrix}
$$
$$
b = [0]
$$

- $\boldsymbol{A}$: $1 \times N$ (chỉ có 1 ràng buộc đẳng thức)

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Tiêu chuẩn Slater cho tối ưu lồi

NẾU: phần trong của tập khả thi (feasible set) không rỗng

THÌ: duality gap = 0 hay **strong duality**

$$\exists(\boldsymbol{w}, b) : f_n(\boldsymbol{w}, b) < 0; \forall n = 1, 2, \cdots, N$$

$\Rightarrow$ **duality gap = 0** hay **strong duality** (lưu ý, dấu $<$, không phải dấu $\leq$)

duality gap = 0 hay **strong duality**:

$$\underbrace{\min_{\boldsymbol{w},b} \underbrace{\max_{\boldsymbol{\alpha}} \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})}_{f_0(\boldsymbol{w},b)}}_{\text{Bài toán gốc}} = \underbrace{\max_{\boldsymbol{\alpha}} \underbrace{\min_{\boldsymbol{w},b} \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\alpha})}_{g(\boldsymbol{\alpha})}}_{\text{Bài toán đối ngẫu}} \tag{3.11}$$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

33 Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Tiêu chuẩn Slater

Ý NGHĨA CỦA TIÊU CHUẨN SLATER và HỆ ĐIỀU KIỆN KKT

Quan hệ hai chiều (cần và đủ):

$(\tilde{w}, \tilde{b}, \tilde{\alpha})$ thoải mãn KKT:

$\Longleftrightarrow$

1. $(\tilde{w}, \tilde{b})$ là nghiệm của bài toán gốc; nghĩa là, $(\tilde{w}, \tilde{b}) \equiv (w^*, b^*)$
2. $(\tilde{\alpha})$ là nghiệm của bài toán đối ngẫu; nghĩa là, $(\tilde{\alpha}) \equiv (\alpha^*)$
3. **duality gap** $= 0$; nghĩa là, $\mathcal{L}(w^*, b^*, \alpha^*) = \frac{1}{2}||w^*||^2$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

34 Tiêu chuẩn Slater

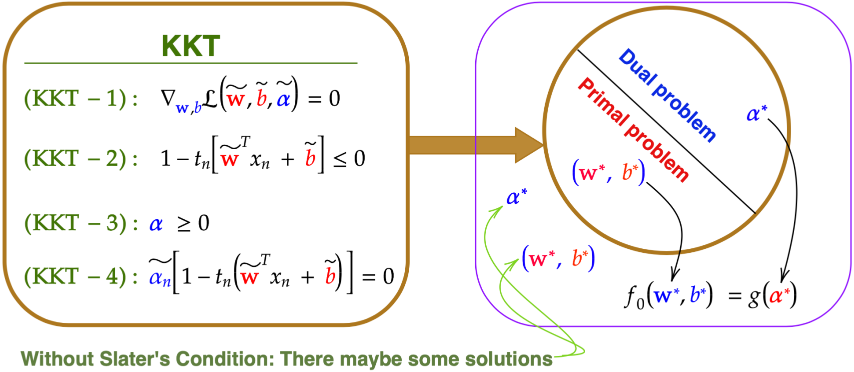
Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Tiêu chuẩn Slater



Hình 3.1: Chỉ riêng hệ điều kiện KKT chưa đủ để bao phủ đầy đủ tập nghiệm của bài toán gốc và đối ngẫu

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
35 Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

Bài toán đối ngẫu

Tiêu chuẩn Slater

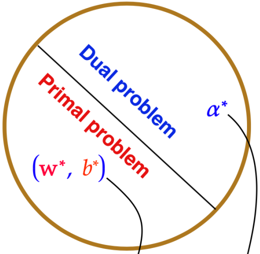
KKT

$$(\text{KKT}-1):\quad \nabla_{\mathbf{w},b}\mathcal{L}\left(\widetilde{\mathbf{w}},\widetilde{b},\widetilde{\alpha}\right)=0$$

$$(\text{KKT}-2):\quad 1-t_n\left[\widetilde{\mathbf{w}}^T x_n+\widetilde{b}\right]\leq 0$$

$$(\text{KKT}-3):\quad \alpha\geq 0$$

$$(\text{KKT}-4):\quad \widetilde{\alpha}_n\left[1-\left(\widetilde{\mathbf{w}}^T x_n+\widetilde{b}\right)\right]=0$$



Slater's condition: $f_0(\mathbf{w}^*, b^*) = g(\alpha^*)$

(strong duality, duality gap = 0)

Chỉ cần tìm nghiệm của hệ điều kiện KKT

Hình 3.2: Khi bài toán gốc thỏa những điều kiện của **KKT** và **Slater** thì chỉ cần tìm nghiệm của hệ điều kiện **KKT** là đủ

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Tiêu chuẩn Slater

Bài toán gốc và T/C Slater

Bài toán gốc thỏa mãn cả hệ điều kiện KKT (ở trên) và tiêu chuẩn Slater

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
Tiêu chuẩn Slater 37
Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Tiêu chuẩn Slater

Từ điều kiện bù

1. $\alpha > 0 \Rightarrow w^T x + b = 1$: các điểm **support** có α tương ứng > 0
2. $w^T x + b > 1 \Rightarrow \alpha = 0$: các điểm không là **support** có α tương ứng $= 0$

Gọi $\mathcal{S}$: tập chứa các chỉ số của các vectơ hỗ trợ

Support Vector Machine
Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
38 Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Công thức dự báo

- Từ C.T (3.2): $\Rightarrow$ véctơ pháp tuyến của siêu phẳng phân lớp bằng tổ hợp tuyến tính của các véctơ hỗ trợ

$$\mathbf{w} = \sum_{s \in \mathcal{S}} \alpha_s t_s \mathbf{x}_s \tag{3.12}$$

- Trong ứng dụng[1], cần xét dấu của hàm phân lớp $y(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$

[1] không cần đến $\mathbf{w}$ thường xuyên, chỉ cần dấu của $y(\mathbf{x})$

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

39 Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu
Công thức dự báo

$$y(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b$$
$$= \sum_{n \in \mathcal{S}} \alpha_n t_n \mathbf{x}_n^T \mathbf{x} + b \quad (3.13)$$

$$y(\mathbf{x}_m) = \sum_{n \in \mathcal{S}} \alpha_n t_n \mathbf{x}_n^T \mathbf{x}_m + b$$

với, $\mathbf{x}_m$: là một véctơ hỗ trợ

$$\Rightarrow b = t_m - \sum_{n \in \mathcal{S}} \alpha_n t_n \mathbf{x}_n^T \mathbf{x}_m$$

$$\Rightarrow b = \frac{1}{N_{\mathcal{S}}} \sum_{m \in \mathcal{S}} \left(t_m - \sum_{n \in \mathcal{S}} \alpha_n t_n \mathbf{x}_n^T \mathbf{x}_m \right) \quad (3.14)$$

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Nguyên tắc
Lagrangian
Xây dựng bài toán
Hàm đối ngẫu
Cực đại hàm đối ngẫu
40 Tiêu chuẩn Slater
Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Công thức dự báo

$$y(\boldsymbol{x}) = \sum_{n \in \mathcal{S}} \alpha_n t_n \boldsymbol{x}_n^T \boldsymbol{x} + \frac{1}{N_{\mathcal{S}}} \sum_{m \in \mathcal{S}} \left(t_m - \sum_{n \in \mathcal{S}} \alpha_n t_n \boldsymbol{x}_n^T \boldsymbol{x}_m \right) \tag{3.15}$$

Vectorization: quy ước

- $\boldsymbol{X}_s$: kích thước $N_s \times (M-1)$, ma trận chứa N_s véctơ hỗ trợ
- $\boldsymbol{X}_b$: kích thước $N_s \times (M-1)$, ma trận chứa N_b véctơ dữ liệu cần dự báo nhãn
- $\boldsymbol{t}_s$: kích thước $N_s \times 1$, véctơ chứa nhãn của N_s véctơ hỗ trợ
- $\boldsymbol{\alpha}_s$: kích thước $N_s \times 1$, véctơ chứa N_s giá trị biến đối ngẫu của các véctơ hỗ trợ
- $\boldsymbol{a}_s = \boldsymbol{\alpha}_s \boldsymbol{t}_s$: nhân phần tử với nhau

ltsach@hcmut.edu.vn
Lê Thành Sách

Bài toán đối ngẫu

Vectorization: Công thức dự báo

- $\boldsymbol{K}_{ss} = \boldsymbol{X}_s \boldsymbol{X}_s^T$
- $\boldsymbol{K}_{bs} = \boldsymbol{X}_b \boldsymbol{X}_s^T$[1]
- $\mathbf{1}$: véctơ cột chứa N_S số 1

$$b = \frac{1}{N_S}\left(\boldsymbol{t}_s - \boldsymbol{K}_{ss}\boldsymbol{a}_s\right)^T \mathbf{1} \tag{3.16}$$

$$\boldsymbol{y} = \boldsymbol{K}_{bs}\boldsymbol{a}_s + b \tag{3.17}$$

$$\text{nhãn} = \mathrm{sign}(\boldsymbol{y}) \tag{3.18}$$

Lưu ý: cả huấn luyện, kiểm thử, và kiểm tra đều được tính thông qua tích vô hướng giữa các điểm dữ liệu.

[1] b: viết tắt của "batch"

Support Vector Machine

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

Nguyên tắc

Lagrangian

Xây dựng bài toán

Hàm đối ngẫu

Cực đại hàm đối ngẫu

42 Tiêu chuẩn Slater

Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Phương pháp mở rộng

Mục lục

Bài toán gốc: ôn lại

Kiến thức liên quan

Bài toán đối ngẫu

43 Phương pháp mở rộng

Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Phương pháp mở rộng

Các trường hợp mở rộng từ hai lớp khả tách tuyến tính

1. Trường hợp hai lớp, **không khả tách tuyến tính**
2. Trường hợp hai lớp, **có biên giới phi tuyến**
3. Trường hợp **nhiều lớp**

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
44 Phương pháp mở rộng
Tổng kết

ltsach@hcmut.edu.vn
Lê Thành Sách

Tổng kết

Mục lục
Bài toán gốc: ôn lại
Kiến thức liên quan
Bài toán đối ngẫu
Phương pháp mở rộng
45 Tổng kết
Tổng kết
Câu hỏi

ltsach@hcmut.edu.vn
Lê Thành Sách

Tổng kết

Bài toán gốc

$$
\begin{aligned}
\boldsymbol{w}^*, b^* &= \underset{\boldsymbol{w}, b}{\operatorname{argmin}} \frac{1}{2}||\boldsymbol{w}||^2 \\
\text{s.t:} \quad & t_n(\boldsymbol{w}^T \boldsymbol{x}_n + b) \geq 1 \\
& \forall n = 1, 2, \ldots, N
\end{aligned}
$$

- Số biến: M
- Số ràng buộc: N
- Dùng khi: $M << N$

Bài toán đối ngẫu

$$
\begin{aligned}
\boldsymbol{\alpha}^* &= \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{K} \boldsymbol{\alpha} + \boldsymbol{p}^T \boldsymbol{\alpha} \\
\text{s.t:} \quad & \boldsymbol{G}\boldsymbol{\alpha} \leq \boldsymbol{h} \\
& \boldsymbol{A}\boldsymbol{\alpha} = b
\end{aligned}
$$

- Số biến: N
- Số ràng buộc: $N + 1$
- Dùng khi: $M > N$

ltsach@hcmut.edu.vn
Lê Thành Sách

Câu hỏi

1. Chứng minh rằng, bài toán gốc thỏa mãn tiêu chuẩn của Slater.
2. Lập trình tạo bộ phân loại SVM, chỉ dùng CVXOPT và numpy.
3. Tại sao nói rằng hàm đối ngẫu là một hàm concave, bất kể rằng bài toán gốc có lồi hay không?
4. Nếu gọi $(\boldsymbol{w}^*, b^*)$ và $\boldsymbol{\alpha}^*$ tương ứng là nghiệm của bài toán gốc và bài toán đối ngẫu. Chứng minh rằng, $\mathcal{L}(\boldsymbol{w}^*, b^*, \boldsymbol{\alpha}^*) = \frac{1}{2}||\boldsymbol{w}^*||^2$
5. Nếu gọi δ là độ rộng lề của bài toán gốc, chứng minh rằng $\delta = \sqrt{2\mathcal{L}(\boldsymbol{w}^*, b^*, \boldsymbol{\alpha}^*)}$
6. Cũng chứng minh rằng, $\delta^2 = \frac{1}{\sum_{n \in \mathcal{S}} \alpha_n}$; với $\mathcal{S}$ là tập chỉ số của các véctơ hỗ trợ.
7. Chưa kể đến độ chính xác của mô hình, chúng ta kỳ vọng một $\mathcal{S}$ có nhiều hay ít phần tử?

ltsach@hcmut.edu.vn
Lê Thành Sách