Ch 9.2: Support Vector Classifier

Lecture 27 - CMSE 381

Prof. Elizabeth Munch. Substitute: Elena Wang

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Wed, Nov 13, 2024

Announcements

Last time:

- 9.1 Maximal Margin Classifier

This lecture:

- 9.2 Support Vector Classifier

Announcements:

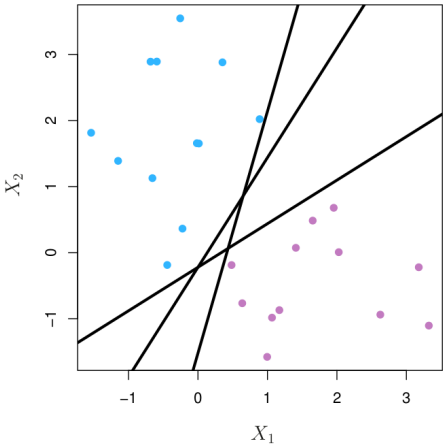
- HW #7 due Sunday 11/17

Lec #	Date			Reading	HW
21	Mon	10/28	Polynomial & Step Functions	7.1,7.2	
22	Wed	10/30	Step Functions; Basis functions; Start Splines	7.2 - 7.4	
23	Fri	11/1	Regression Splines	7.4	HW #6 Due Sun 11/3
24	Mon	11/4	Decision Trees	8.1	
25	Wed	11/6	Class Cancelled (Dr Munch out of town)		
26	Fri	11/8	Random Forests	8.2.1, 8.2.2	HW #7 Due Sun 11/10
27	Mon	11/11	Maximal Margin Classifier	9.1	
28	Wed	11/13	SVC	9.2	
29	Fri	11/15	SVM	9.3, 9.4	HW #8 Due Sun 11/17
30	Mon	11/18	Single layer NN	10.1	
31	Wed	11/20	Multi Layer NN	10.2	
32	Fri	11/22	CNN	10.3	HW #11 Due Sun 11/24
33	Mon	11/25	TBD: Unsupervised learning/clustering	12.1, 12.4?	
	Wed	11/27	Virtual: Project office hours		
	Fri	11/29	No class - Thanksgiving		
	Mon	12/2	Review		
	Wed	12/4	**Midterm #3**		
	Fri	12/6	No class - EGR Design Day		Project due

Section 1

Last time

Separating Hyperplane



Require that for every data point:

$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} > 0 \text{ if } y_i = 1$$
$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} < 0 \text{ if } y_i = -1$$

Equivalently

Require that for every data point

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) > 0$$

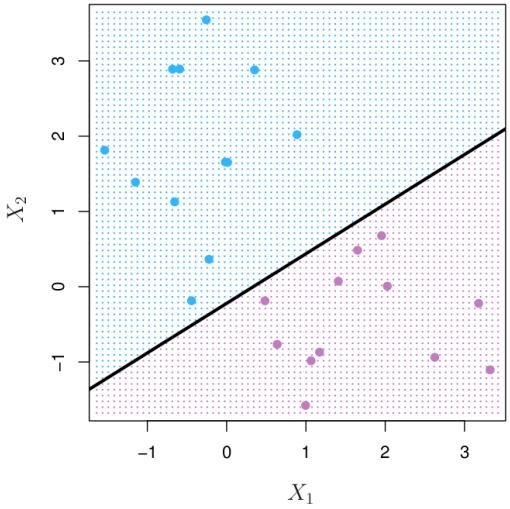
Separating hyperplane becomes a classifier

If you have a separating hyperplane:

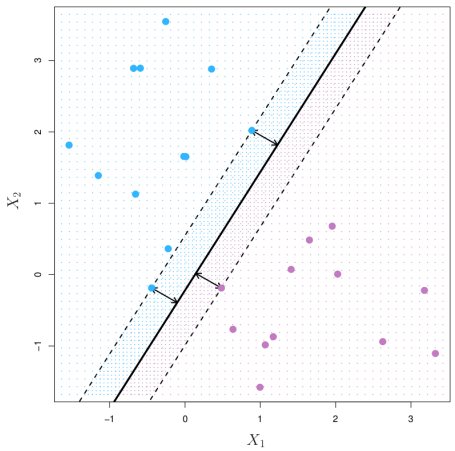
- Check

$$f(x^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \cdots + \beta_p x_p^*$$

- If positive, assign $\hat{y} = 1$
- If negative, assign $\hat{y} = -1$



Maximal margin classifier



- For a hyperplane, the *margin* is the smallest distance from any data point to the hyperplane.
- Observations that are closest are called *support vectors*.
- The *maximal margin hyperplane* is the hyperplane with the largest margin
- The classifier built from this hyperplane is the *maximal margin classifier*.

Mathematical Formulation

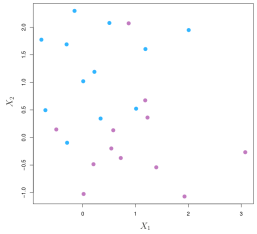
$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{maximize}} \; M$$

$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

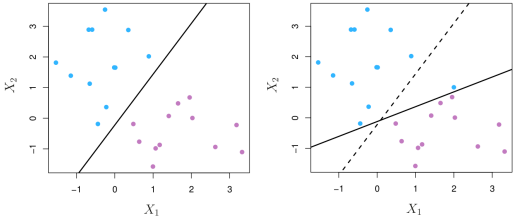
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M \;\; \forall \; i = 1, \ldots, n$$

Problems

Might be no separating hyperplane



Sensitivity to new points

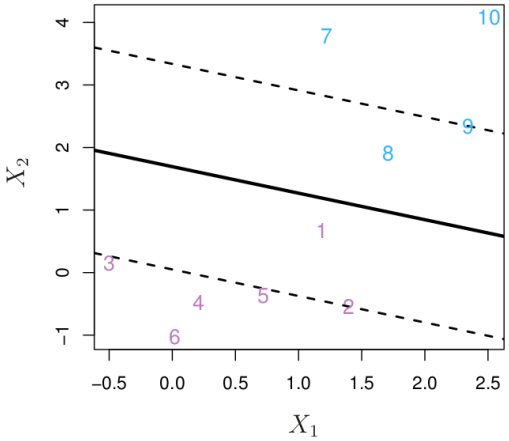


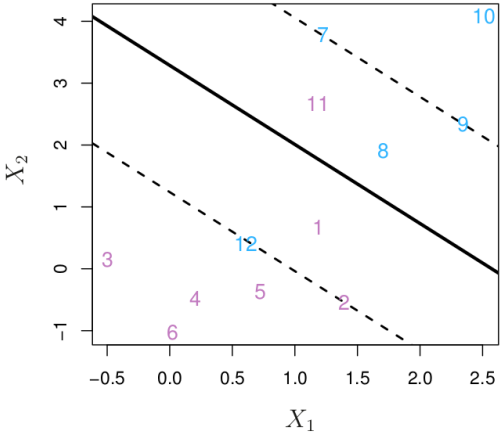
Section 2

Support Vector Classifier

Basic idea

Soft margin





Mathematical Formulation of SVC

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,\,M}{\text{maximize}} \quad M$$

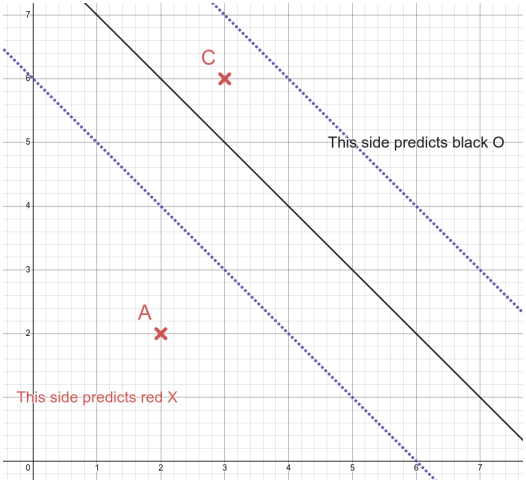
$$\text{subject to} \quad \sum_{j=1}^{p} \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \epsilon_i),$$

$$\epsilon_i \geq 0, \quad \sum_{i=1}^{n} \epsilon_i \leq C,$$

Find positive ε's that will satisfy this

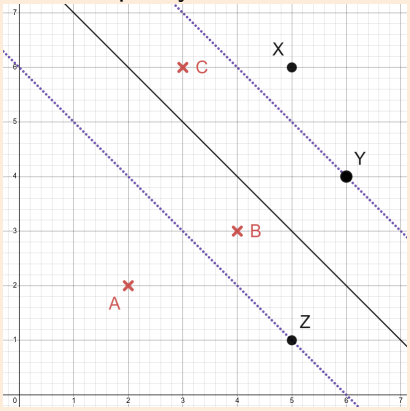
Fix $M = \sqrt{2}$ $\quad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$



What is ε?

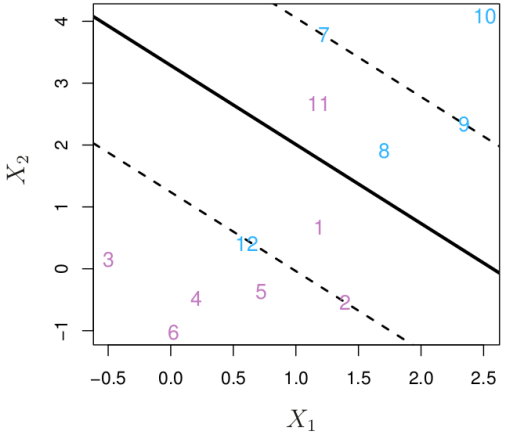
Fix $M = \sqrt{2}$ $\qquad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$

Fill in the table so that the inequality is satisfied.



Point	Left Side	ε_i	$M(1-\varepsilon_i)$
A	$2\sqrt{2}$	0	$\sqrt{2}$
B	$\frac{\sqrt{2}}{2}$		
C	$-\frac{\sqrt{2}}{2}$	1.5	$-\frac{\sqrt{2}}{2}$
X	$\frac{3\sqrt{2}}{2}$		
Y	$\sqrt{2}$		
Z	$-\sqrt{2}$		

What is ε?



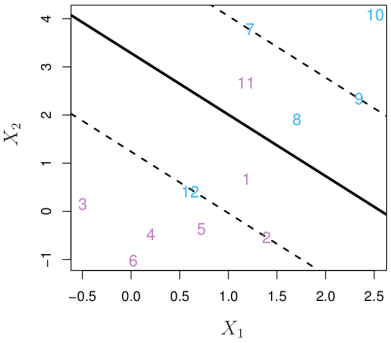
What is C?

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,\, M}{\text{maximize}} \quad M$$

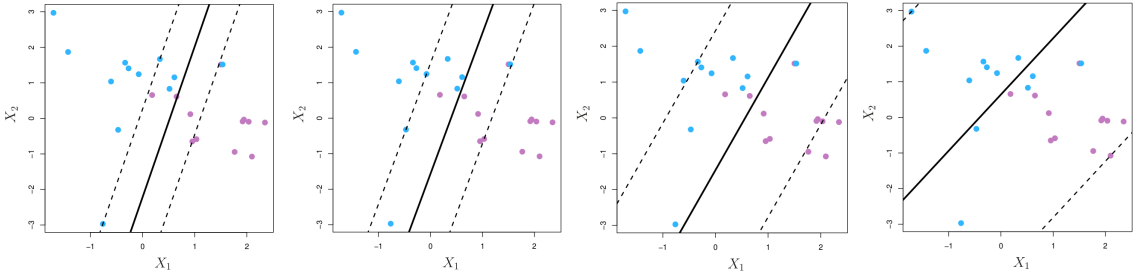
$$\text{subject to} \quad \sum_{j=1}^{p} \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \epsilon_i),$$

$$\epsilon_i \geq 0, \quad \sum_{i=1}^{n} \epsilon_i \leq C,$$

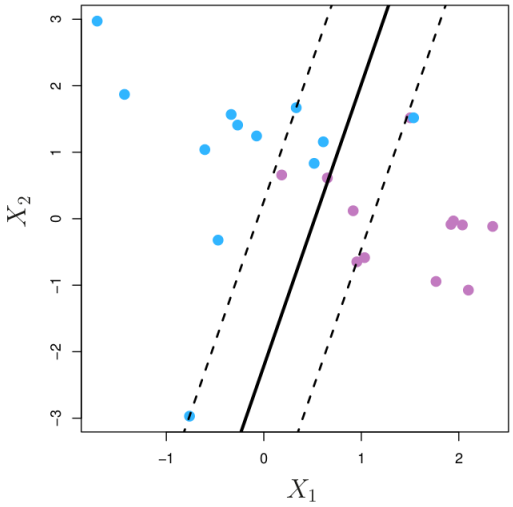


Examples messing with C



Increasing $C \rightarrow$

What affects the hyperplane?



Coding

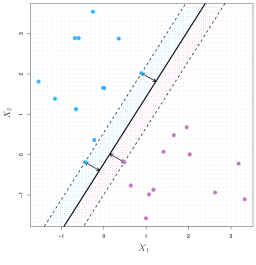
TL;DR

Maximal Margin Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{maximize}} \; M$$

$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M \;\; \forall \, i = 1, \ldots, n$$



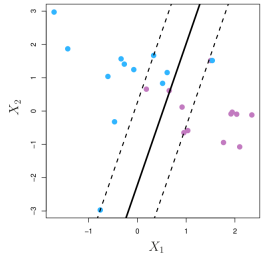
Support Vector Classifier

$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,\, M}{\text{maximize}} \; M$$

$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \epsilon_i),$$

$$\epsilon_i \geq 0, \;\; \sum_{i=1}^{n} \epsilon_i \leq C,$$



Next time

Lec #	Date			Reading	HW
21	Mon	10/28	Polynomial & Step Functions	7.1,7.2	
22	Wed	10/30	Step Functions; Basis functions; Start Splines	7.2 - 7.4	
23	Fri	11/1	Regression Splines	7.4	HW #6 Due Sun 11/3
24	Mon	11/4	Decision Trees	8.1	
25	Wed	11/6	Class Cancelled (Dr Munch out of town)		
26	Fri	11/8	Random Forests	8.2.1, 8.2.2	HW #7 Due Sun 11/10
27	Mon	11/11	Maximal Margin Classifier	9.1	
28	Wed	11/13	SVC	9.2	
29	Fri	11/15	SVM	9.3, 9.4	HW #8 Due Sun 11/17
30	Mon	11/18	Single layer NN	10.1	
31	Wed	11/20	Multi Layer NN	10.2	
32	Fri	11/22	CNN	10.3	HW #11 Due Sun 11/24
33	Mon	11/25	TBD: Unsupervised learning/clustering	12.1, 12.4?	
	Wed	11/27	Virtual: Project office hours		
	Fri	11/29	No class - Thanksgiving		
	Mon	12/2	Review		
	Wed	12/4	**Midterm #3**		
	Fri	12/6	No class - EGR Design Day		Project due