

Ch 9.2: Support Vector Classifier

Lecture 27 - CMSE 381

Prof. Guanqun Cao

Michigan State University

::

Dept of Computational Mathematics, Science & Engineering

Fri, Nov 7, 2025

Announcements

Last time:

- 9.1 Maximal Margin Classifier

This lecture:

- 9.2 Support Vector Classifier

Announcements:

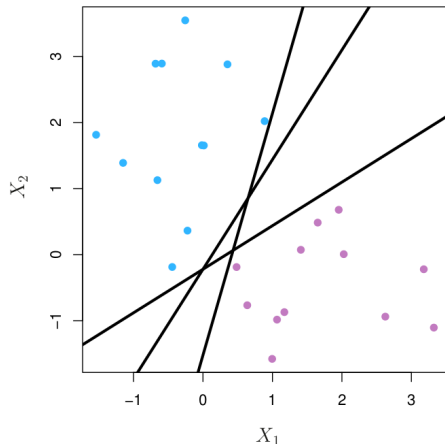
- HW #7 due Sunday

	F	10/17	Review		
	M	10/20	Fall Break		
	W	10/22	Midterm #2		
21	F	10/24	Polynomial & Step Functions	7.1-7.2	HW #5 Due Sun 10/28
22	M	10/27	Step Functions; Basis functions; Start Splines	7.2-7.4	
23	W	10/29	Regression Splines	7.4	
24	F	10/31	Decision Trees	8.1	HW #6 Due Sun 11/2
25	M	11/3	Random Forests	8.2.1, 8.2.2	
26	W	11/5	Maximal Margin Classifier	9.1	
27	F	11/7	SVC	9.2	HW #7 Due Sun 11/9
28	M	11/10	SVM	9.3, 9.4	
29	W	11/12	Single Layer NN	10.1	
30	F	11/13	Multi Layer NN	10.2	HW #8 Due Sun 11/16
31	M	11/17	CNN	10.3	
32	W	11/19	Unsupervised learning / clustering	12.1, 12.4	
33	F	11/21	Virtual: Project Office Hours		HW #9 Due Sun 11/23
	M	11/24	Review		
	W	11/26	Midterm #3		
	F	11/28	Thanksgiving		
	M	12/1	Virtual: Project Office Hours		
	W	12/3	Virtual: Project Office Hours		
	F	12/5			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

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

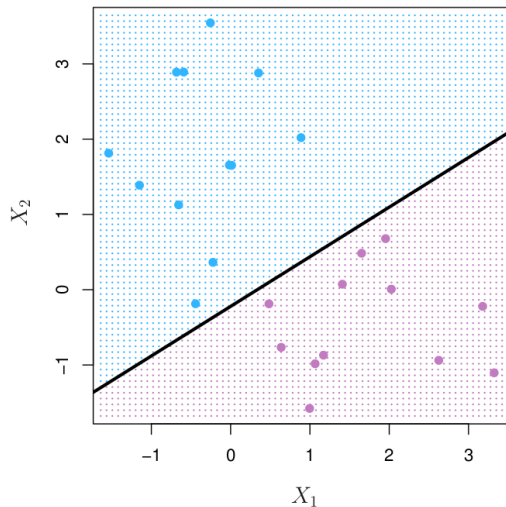
Separating hyperplane becomes a classifier

If you have a separating hyperplane:

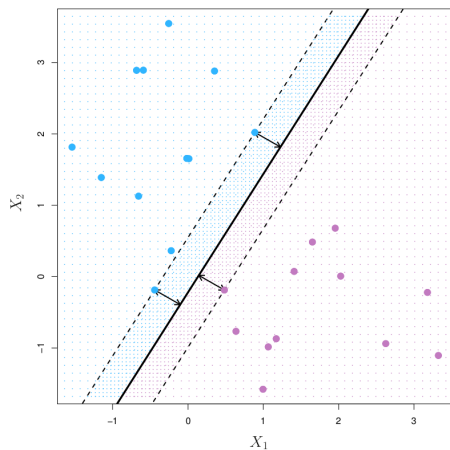
- Check

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

- 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*.

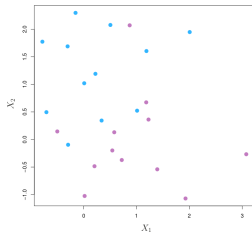
Mathematical Formulation

$$\underset{\beta_0, \beta_1, \dots, \beta_p, M}{\text{maximize}} \quad M$$

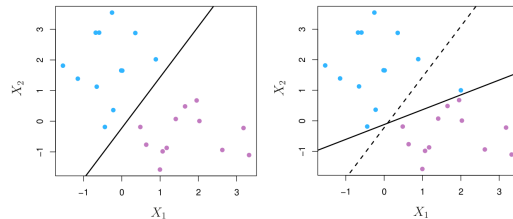
$$\text{subject to} \quad \sum_{j=1}^p \beta_j^2 = 1,$$

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

Might be no separating hyperplane



Sensitivity to new points

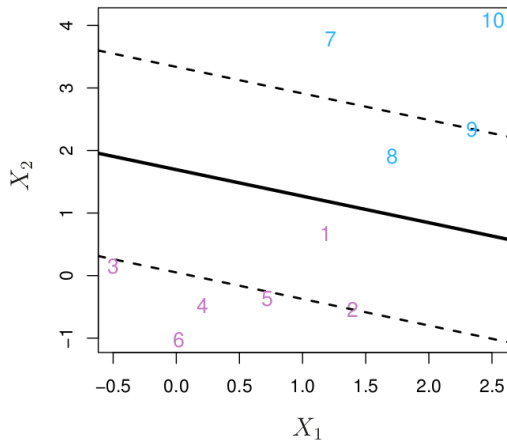


Section 2

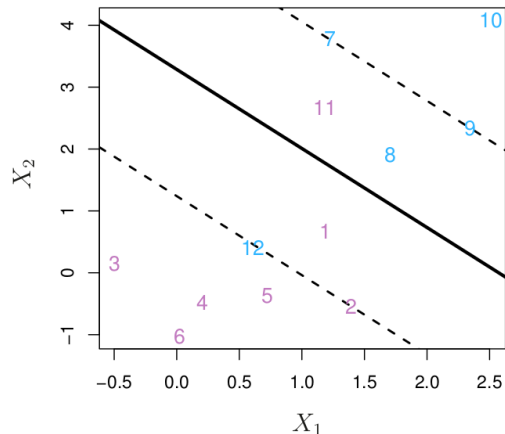
Support Vector Classifier

Basic idea

Soft margin



Some points on wrong side of margin



*Some points on wrong side of hyperplane
(Misclassified)*

Mathematical Formulation of SVC

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

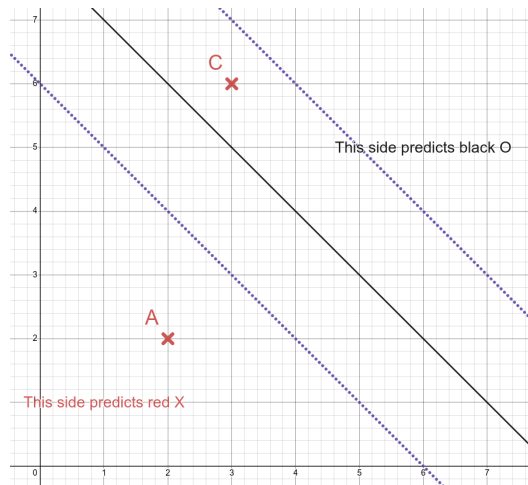
$$\text{subject to} \quad \sum_{j=1}^p \beta_j^2 = 1,$$

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

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

Find positive ε 's that will satisfy this

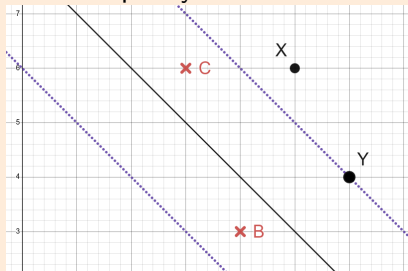
$$\text{Fix } M = \sqrt{2} \quad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$$



What is ε ?

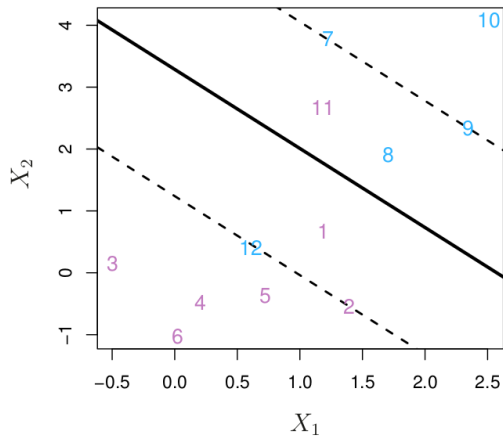
Fix $M = \sqrt{2}$ $y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$

Fill in the table so that the inequality is satisfied.



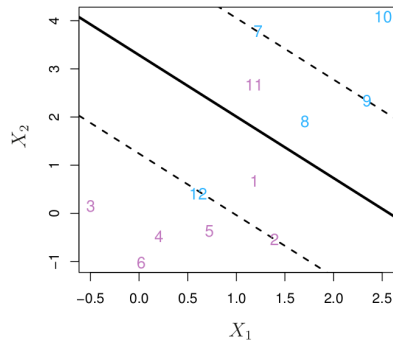
Point	Left Side	ε_i	$M(1 - \varepsilon_i)$
A	$2\sqrt{2}$	0	$\sqrt{2}$
B	$\frac{\sqrt{2}}{2}$	1.5	$-\frac{\sqrt{2}}{2}$
C	$-\frac{\sqrt{2}}{2}$		
X	$\frac{3\sqrt{2}}{2}$		
Y	$\sqrt{2}$		
Z	$-\sqrt{2}$		

What is ε ?

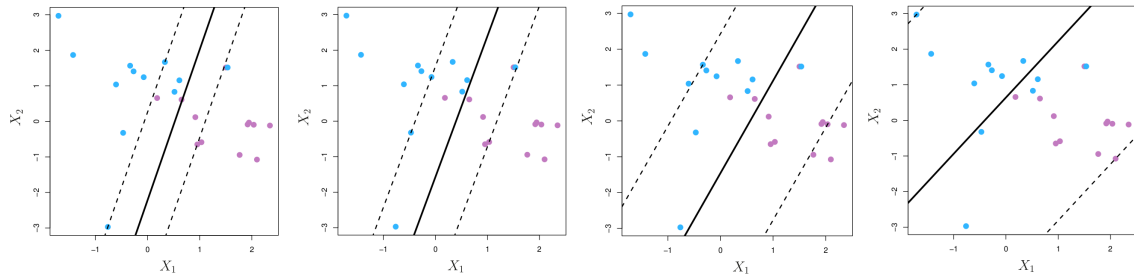


What is C ?

$$\begin{aligned} & \underset{\beta_0, \beta_1, \dots, \beta_p, \epsilon_1, \dots, \epsilon_n, M}{\text{maximize}} && M \\ & \text{subject to} && \sum_{j=1}^p \beta_j^2 = 1, \\ & && y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) \geq M(1 - \epsilon_i), \\ & && \epsilon_i \geq 0, \quad \sum_{i=1}^n \epsilon_i \leq C, \end{aligned}$$

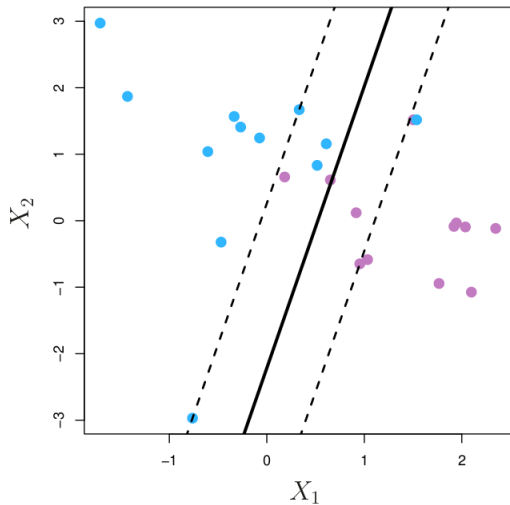


Examples messing with C



Increasing $C \rightarrow$

What affects the hyperplane?



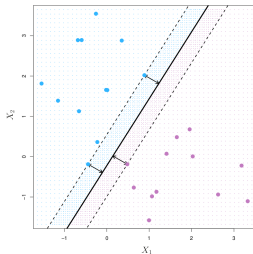
Coding

Maximal Margin Classifier

$$\underset{\beta_0, \beta_1, \dots, \beta_p, M}{\text{maximize}} \quad M$$

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

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



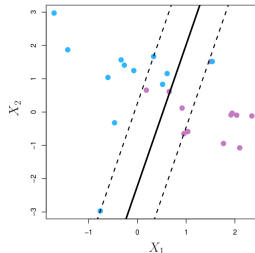
Support Vector Classifier

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

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

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

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



Next time

	F	10/17	Review		
	M	10/20	Fall Break		
	W	10/22	Midterm #2		
21	F	10/24	Polynomial & Step Functions	7.1-7.2	HW #5 Due Sun 10/28
22	M	10/27	Step Functions; Basis functions; Start Splines	7.2-7.4	
23	W	10/29	Regression Splines	7.4	
24	F	10/31	Decision Trees	8.1	HW #6 Due Sun 11/2
25	M	11/3	Random Forests	8.2.1, 8.2.2	
26	W	11/5	Maximal Margin Classifier	9.1	
27	F	11/7	SVC	9.2	HW #7 Due Sun 11/9
28	M	11/10	SVM	9.3, 9.4	
29	W	11/12	Single Layer NN	10.1	
30	F	11/13	Multi Layer NN	10.2	HW #8 Due Sun 11/16
31	M	11/17	CNN	10.3	
32	W	11/19	Unsupervised learning / clustering	12.1, 12.4	
33	F	11/21	Virtual: Project Office Hours		HW #9 Due Sun 11/23
	M	11/24	Review		
	W	11/26	Midterm #3		
	F	11/28	Thanksgiving		
	M	12/1	Virtual: Project Office Hours		
	W	12/3	Virtual: Project Office Hours		
	F	12/5			Project Due

Q of the Day: what is ϵ in $\dots > M(1 - \epsilon_i)$ called?