Ch 9.2: Support Vector Classifier

Lecture 26 - CMSE 381

Prof. Elizabeth Munch

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

Fri, Nov 18, 2022

Announcements

Last time:

- 9.1 Maximal Margin Classifier

This lecture:

- 9.2 Support Vector Classifier

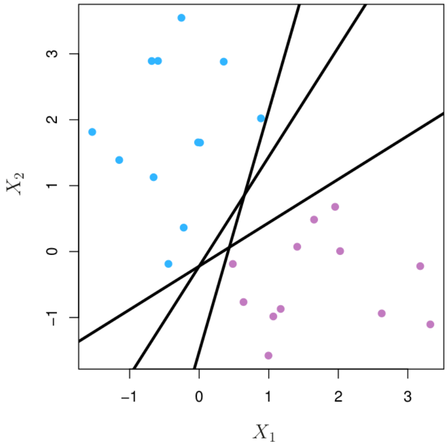
Announcements:

- HW #8 due tonight

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

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

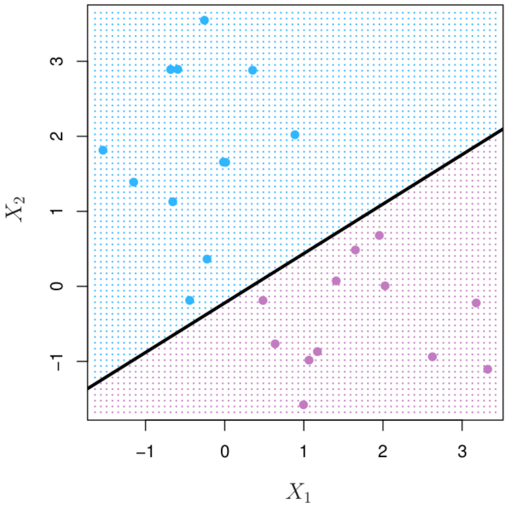
Separating hyperplane becomes a classifier

If you have a separating hyperplane:

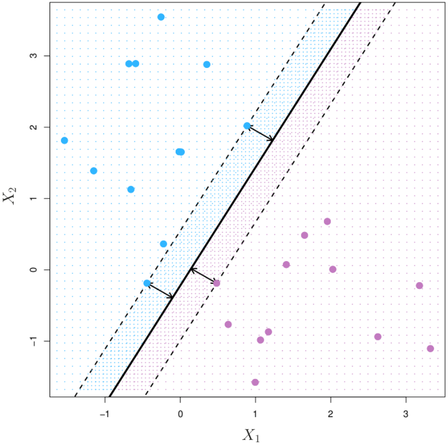
- Check

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

- 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

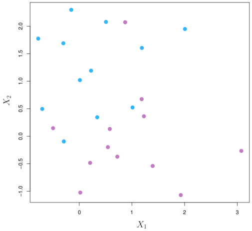
$$\underset{\beta_0,\beta_1,\ldots,\beta_p,M}{\text{maximize}}\ M$$

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

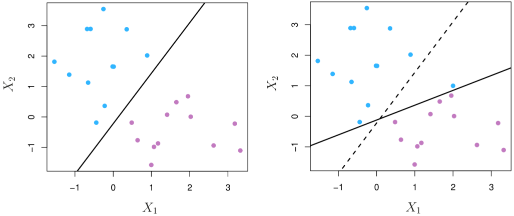
$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M \ \ \forall\, i = 1, \ldots, n$$

Problems

Might be no separating hyperplane



Sensitivity to new points

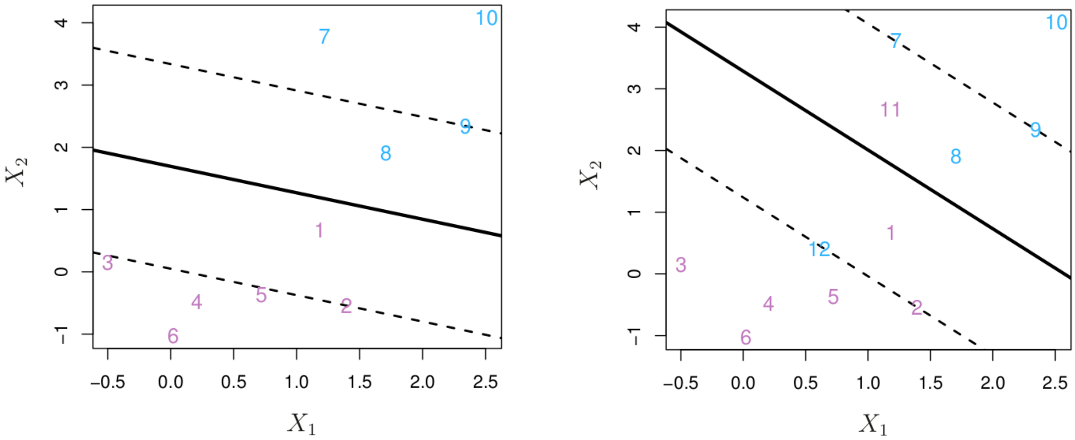


Section 2

Support Vector Classifier

Basic idea

Soft margin



10
7
9
8
1
3
4
5
2
6
−1
0
1
2
3
4
X_2
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
X_1
10
7
11
9
8
1
12
3
5
4
2
6
X_2
X_1

Mathematical Formulation of SVC

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

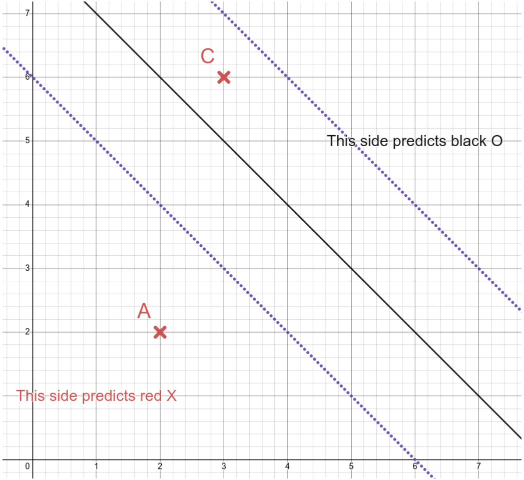
$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

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

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

Find positive ε's that will satisfy this

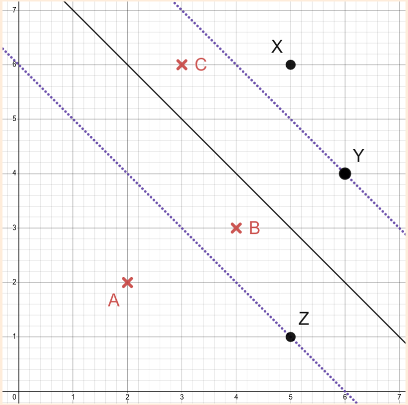
Fix $M = \sqrt{2}$ $\quad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$



What is ε?

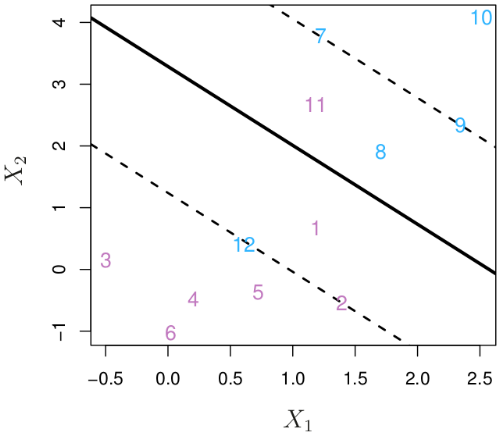
$$\text{Fix } M = \sqrt{2} \qquad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$$

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?



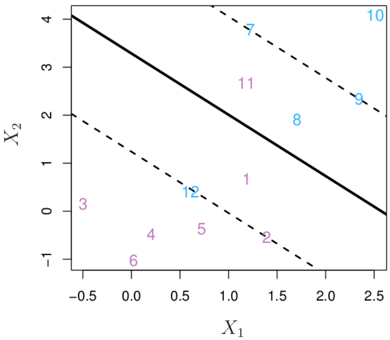
What is C?

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

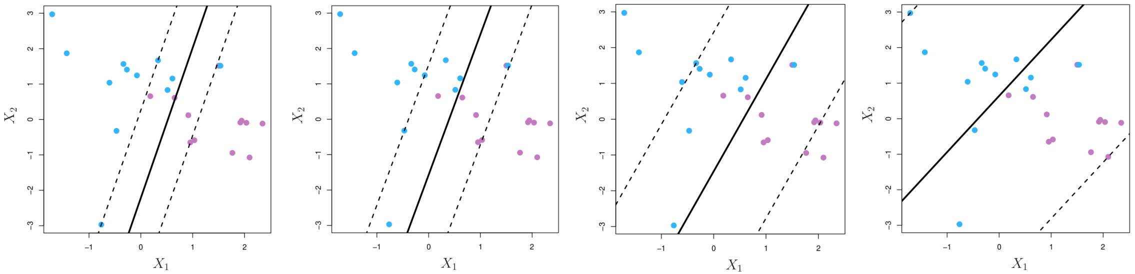
$$\text{subject to} \quad \sum_{j=1}^{p} \beta_j^2 = 1,$$

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

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

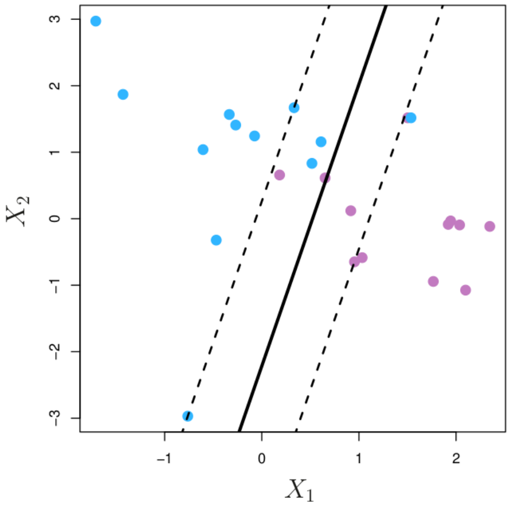


Examples messing with C



Increasing C $\rightarrow$

What affects the hyperplane?



Coding

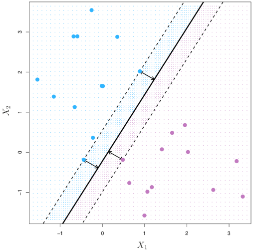
TL;DR

Maximal Margin Classifier

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

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

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



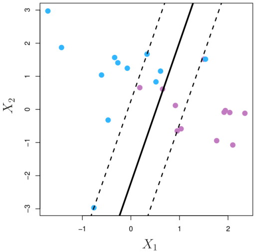
Support Vector Classifier

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

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

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

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



Next time

20	F	Nov 4	Polynomial & Step Functions.	7.1,7.2	
21	M	Nov 7	Step Functions	7.2	
22	W	Nov 9	Basis functions, Regression Splines	7.3,7.4	
23	F	Nov 11	Decision Trees	8.1	HW #7 Due
24	M	Nov 14	Random Forests	8.2.1, 8.2.2	
25	W	Nov 16	Maximal Margin Classifier	9.1	
26	F	Nov 18	SVC	9.2	HW #8 Due
27	M	Nov 21	SVM	9.3, 9.4, 9.5	
28	W	Nov 23	Extended virtual office hours		
	F	Nov 25	No class - Thanksgiving		
29	M	Nov 28	Single layer NN	10.1	HW #9 Due
30	W	Nov 30	Multi Layer NN	10.2	
31	F	Dec 2	CNN	10.3	
32	M	Dec 5	Unsupervised Learning & Clustering	12.1, 12.4	HW #10 Due
	W	Dec 7	Review		
	F	Dec 9	Midterm #3	Bring your cheat sheet and a non-internet-connected calculator	