Ch 9.2: Support Vector Classifier
Lecture 28 - CMSE 381

Prof. Elizabeth Munch

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

Wed, Nov 15, 2023

Announcements

Last time:

- 9.1 Maximal Margin Classifier

This lecture:
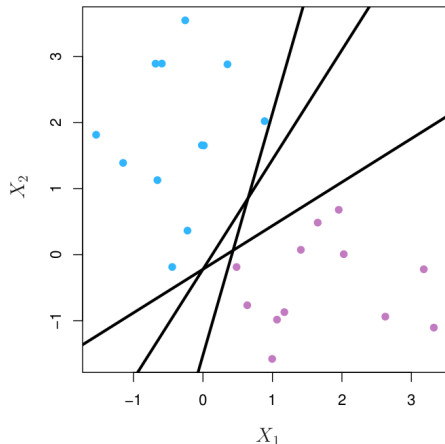
- 9.2 Support Vector Classifier

Announcements:

- HW #7....

Section 1

Last time

Separating Hyperplane

Require that for every data point:

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

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

Equivalently

Require that for every data point

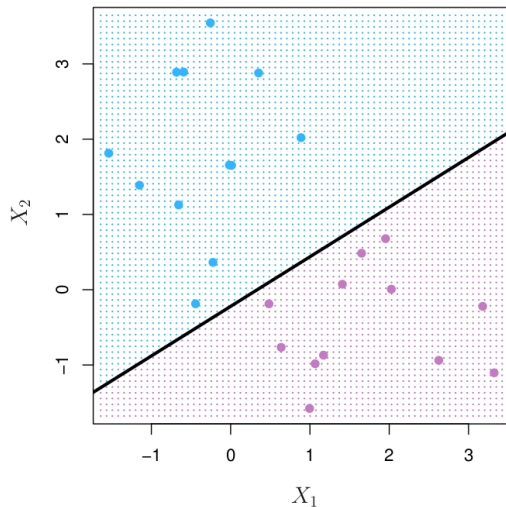$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) > 0$

Separating hyperplane becomes a classifier

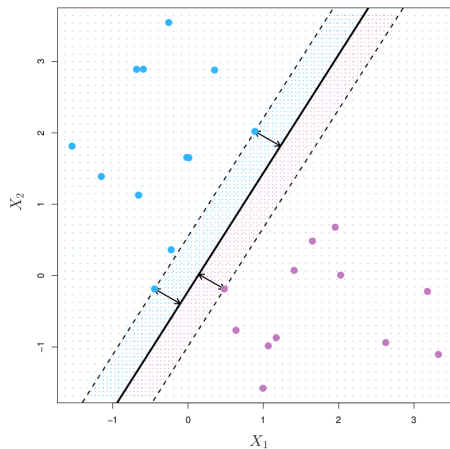If you have a separating hyperplane:

- 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
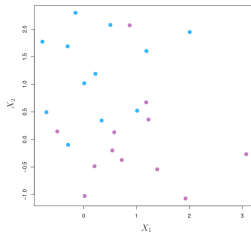- The classifier built from this hyperplane is the *maximal margin classifier*.

Mathematical Formulation

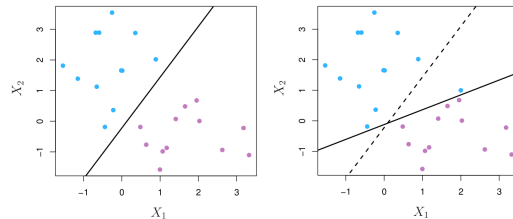$$\underset{\beta_0, \beta_1, \ldots, \beta_p, M}{\text{maximize}} \; M$$

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

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

Problems

Might be no separating hyperplane

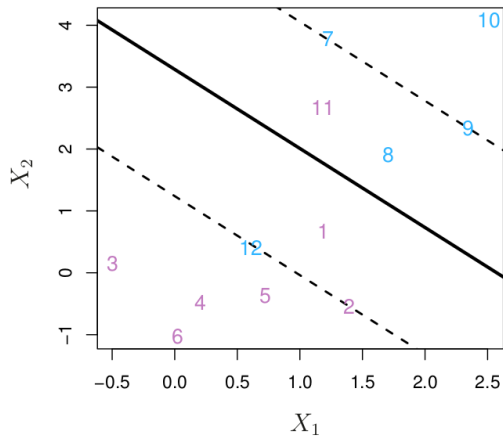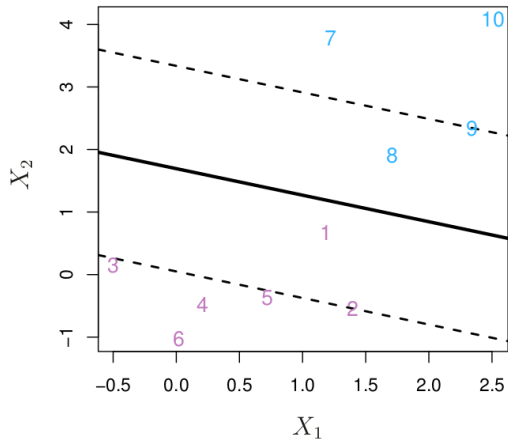
Sensitivity to new points

Section 2

Support Vector Classifier

Basic idea

Soft margin

Mathematical Formulation of SVC

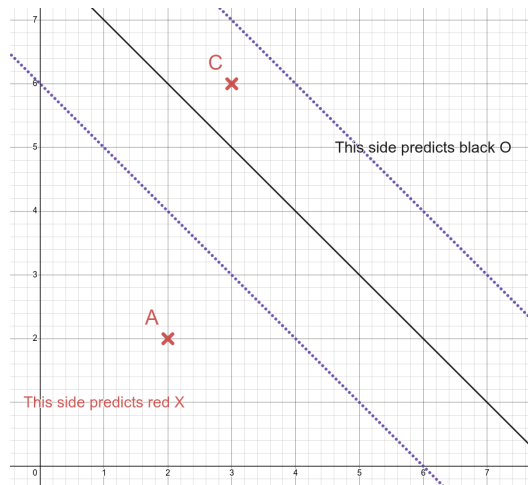$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,\, M}{\text{maximize}} \quad M$$

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

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

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

Find positive ε's that will satisfy this

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

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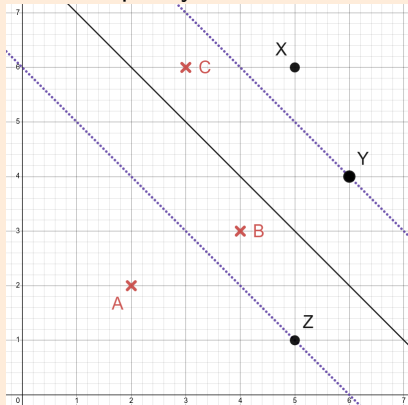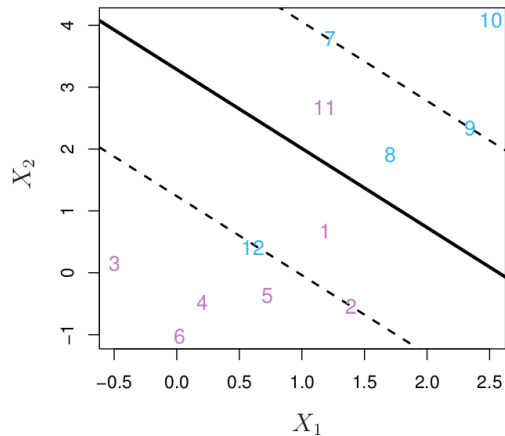
$$\text{Fix } M = \sqrt{2} \qquad y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i)$$

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?

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

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

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

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

Examples messing with C

Increasing $C \rightarrow$

What affects the hyperplane?
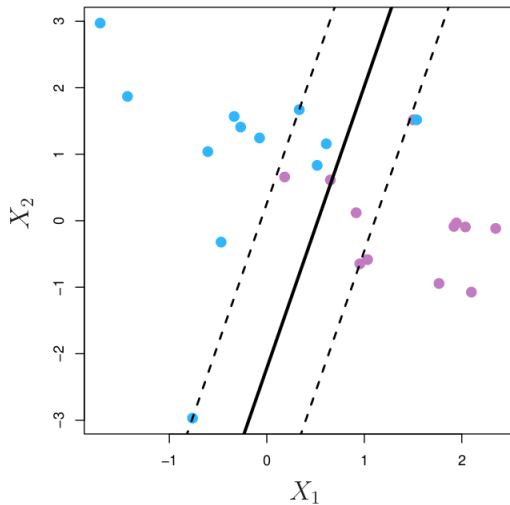
Coding

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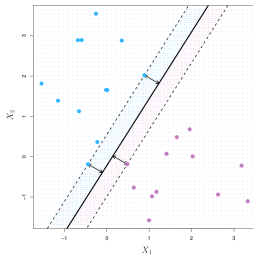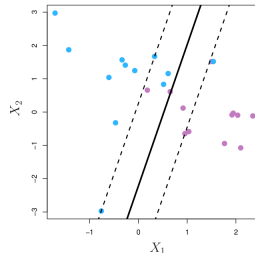
$$\underset{\beta_0,\beta_1,\ldots,\beta_p,\epsilon_1,\ldots,\epsilon_n,M}{\text{maximize}} \; M$$

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

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

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

Next time

Status	Lec #		Date		Reading	Homeworks
		Mon	Oct 23	No class - Fall break		
		Wed	Oct 25	**Midterm #2**		
Done	20	Fri	Oct 27	Dimension Reduction	6.3	
Done	21	Mon	Oct 30	More dimension reduction; High dimensions	6.4	
Done	22	Wed	Nov 1	Polynomial & Step Functions	7.1,7.2	
Pushed	23	Fri	Nov 3	Step Functions; Basis functions; Start Splines	7.2 - 7.4	
	24	Mon	Nov 6	Regression Splines	7.4	~~HW #6 Due~~
	25	Wed	Nov 8	Decision Trees	8.1	HW #6 Due
	26	Fri	Nov 10	Random Forests	8.2.1, 8.2.2	
	27	Mon	Nov 13	Maximal Margin Classifier	9.1	
	28	Wed	Nov 15	SVC	9.2	
	29	Fri	Nov 17	SVM	9.3, 9.4	
	30	Mon	Nov 20	Single layer NN	10.1	
	31	Wed	Nov 22	Virtual: Project office hours		
		Fri	Nov 24	No class - Thanksgiving		
		Mon	Nov 27	**Review**		
		Wed	Nov 29	**Midterm #3**		
	32	Fri	Dec 1	Multi Layer NN	10.2	
	33	Mon	Dec 4	CNN	10.3	
	34	Wed	Dec 6	Unsupervised Learning & Clustering	12.1, 12.4	
	35	Fri	Dec 8	Virtual: Project office hours		Project due