Ch 9.1: Maximum Margin Classifier

Lecture 26 - CMSE 381

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

Mon, Mar 30, 2026

Announcements

Last time:

- Ch 8: Random Forests

This lecture:

- Maximal Margin Classifier
- No jupyter notebook for this class

Announcements:

21	W	3/18	Polynomial & Step Functions	7.1-7.2		
22	F	3/20	Step Functions; Basis functions; Start Splines	7.2-7.4		
23	M	3/23	Regression Splines	7.4		
24	W	3/25	Decision Trees	8.1		Q7
25	F	3/27	Random Forests	8.2.1, 8.2.2	HW #5 Due Sun 3/29	
26	M	3/30	Maximal Margin Classifier	9.1		
27	W	4/1	SVC	9.2		Q8
28	F	4/3	SVM	9.3, 9.4		
29	M	4/6	Single Layer NN	10.1		
30	W	4/8	Multi Layer NN	10.2		Q9
31	F	4/10	CNN	10.3		
32	M	4/13	Unsupervised learning / clustering	12.1, 12.4	HW #6 Due Sun 4/12	Q10
33	W	4/15	Virtual: Project Office Hours			
	F	4/17	***Review***			
	M	4/20	**Midterm #3**			
	W	4/22				
	F	4/24			**Project Due**	

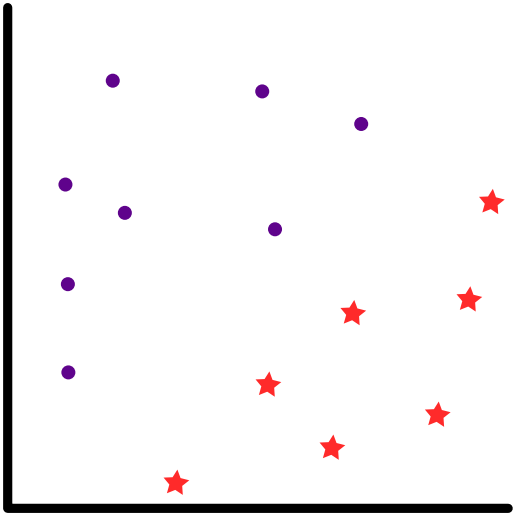
What will you learn today?

- What is hyperplane?
  - How do you mathematically describe this hyperplane?
  - How do you mathematically describe the two sides of the hyperplane?
  - Given the equation of the hyperplane and the coordinates for a point, you should be able to tell which side of the plane the point is on.
- What qualify a hyperplane as a separating hyperplane?
  - You should be able to describe this mathematically using an inequality.
  - You should also be able to determine whether a hyperplane is a separating hyperplane given a graph, or given the equation of the plane and the coordinates and class of a few points.
- How to use a separating hyperplane as a classifier?
- What makes a hyperplane a maximal margin hyperplane?
  - What are its margin and support vectors?
  - Given a graph, you should be able to clearly label the margin and support vectors. You should also be able to infer the size of the margin from reading the graph.
  - You should also be able to describe the optimization problem mathematically.
  - Given the equation of the maximal margin hyperplane and the coordinates of support vectors, you should be able to calculate the size of the margin by hand.

Section 1

Maximal Margin Classifier

The goal

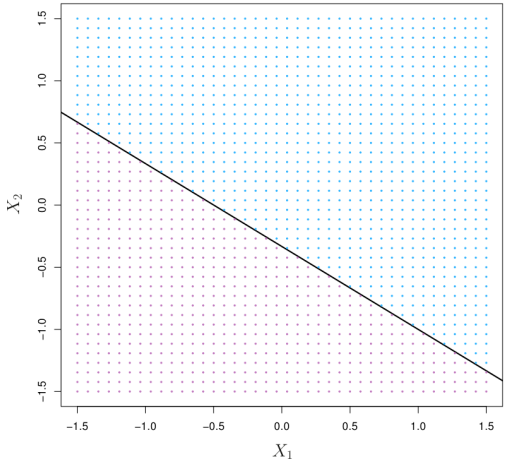


What is a hyperplane?

Mathematical definition of a hyperplane

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p = 0$$

Hyperplane for $p = 2$

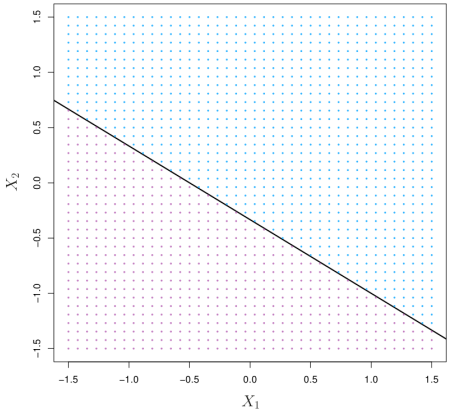


$$1 + 2X_1 + 3X_2 = 0$$

There are two sides to every hyperplane

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p < 0$$

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p > 0$$



Classification Setup

Data matrix:

$$X = \begin{pmatrix} - & x_1^T & - \\ - & x_2^T & - \\ & \vdots & \\ - & x_n^T & - \end{pmatrix}_{n \times p}$$

Observations in one of two classes, $y_i \in \{-1, 1\}$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

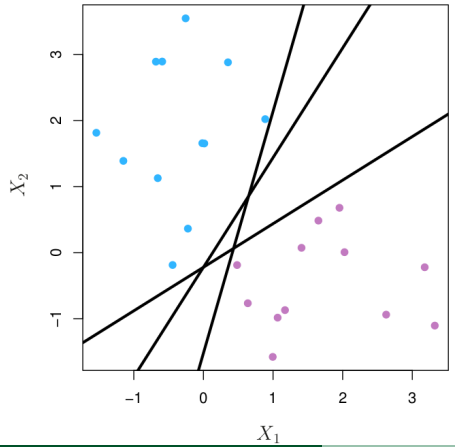
$$x_1 = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1p} \end{pmatrix}, \cdots, x_n = \begin{pmatrix} x_{n1} \\ \vdots \\ x_{np} \end{pmatrix}$$

Separate out a test observation

$$x^* = (x_1^* \ \cdots \ x_p^*)^T$$

Separating Hyperplane

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



Another way to say it

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

For all i:

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

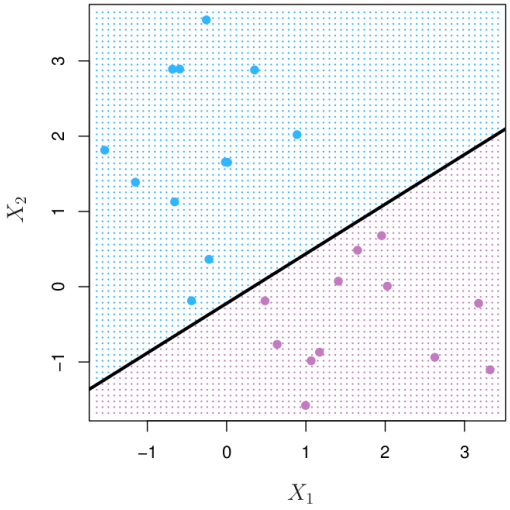
Separating hyperplane becomes a classifier

If you have a separating hyperplane:

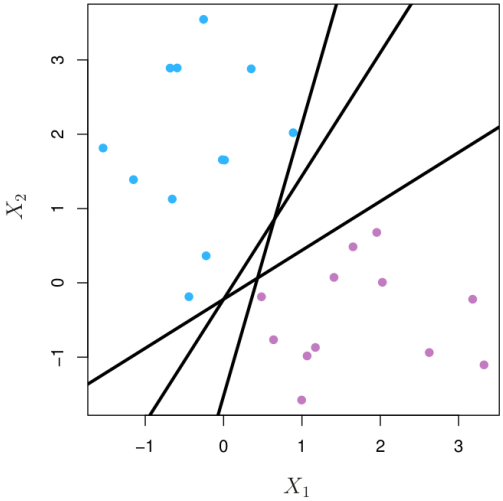
- Check

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

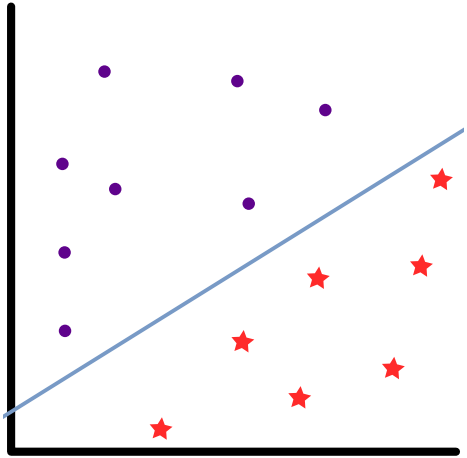
- If positive, assign $\hat{y} = 1$
- If negative, assign $\hat{y} = -1$

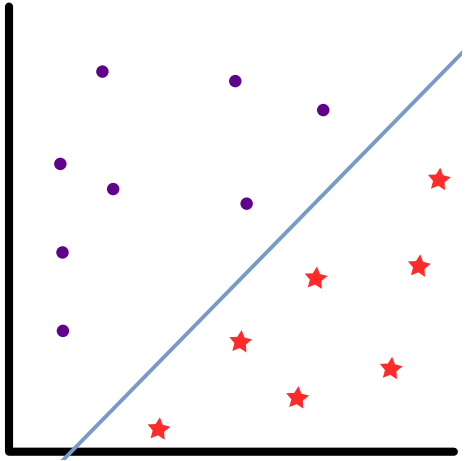


How do we pick?

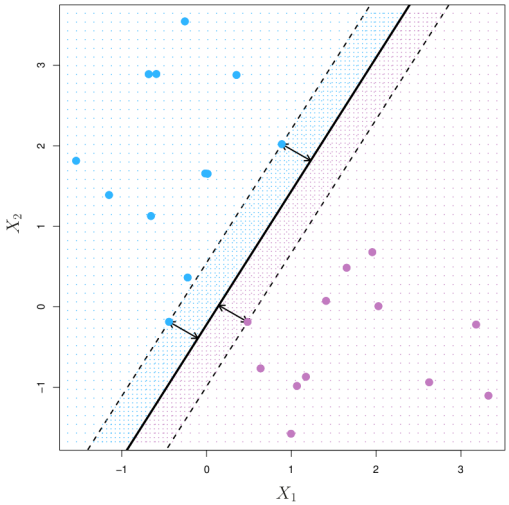


Distance from an observation to a hyperplane

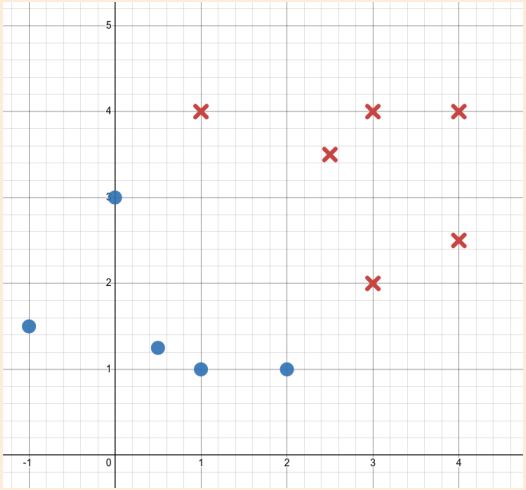




Maximal margin classifier



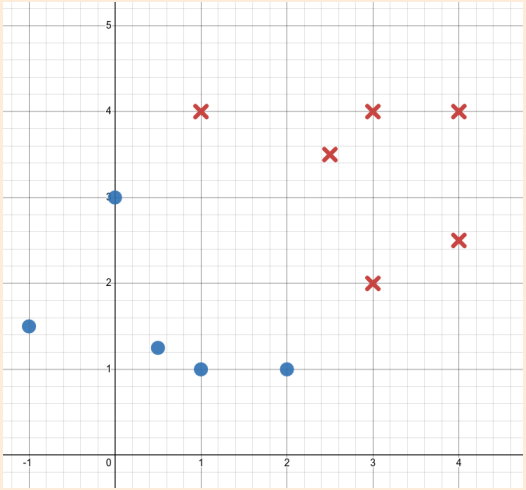
Example



- Sketch the maximal margin hyperplane.
- What is the equation of this line in the form $\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0$?
- Circle the support vectors. What is their distance from the line?

https://www.desmos.com/calculator/tklbommiwz

Extra work space



Respond to PollEv question.

https://www.desmos.com/calculator/tklbommiwz

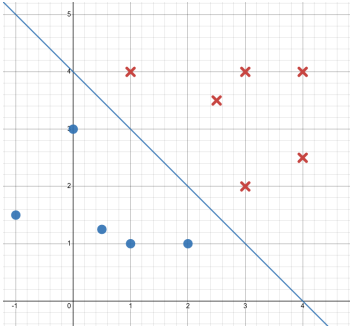
Mathematical Formulation

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

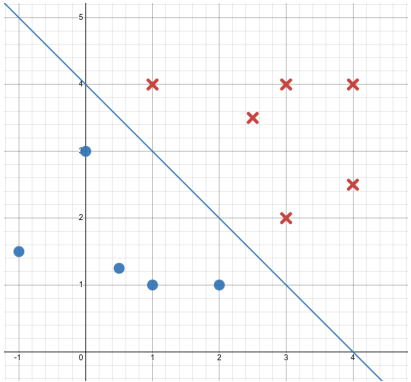
$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

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

First constraint



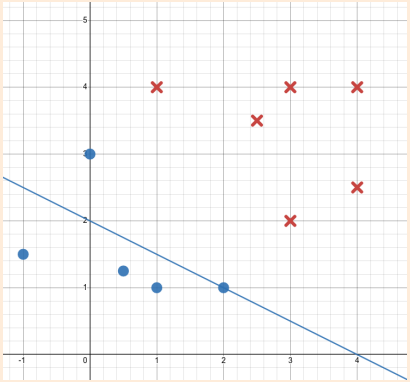
Second constraint



$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}) \geq M$$

- Blue circles: $y_i = -1$
- Red Xs: $y_i = 1$
- $-2\sqrt{2} + \frac{\sqrt{2}}{2} X_1 + \frac{\sqrt{2}}{2} X_1 = 0$

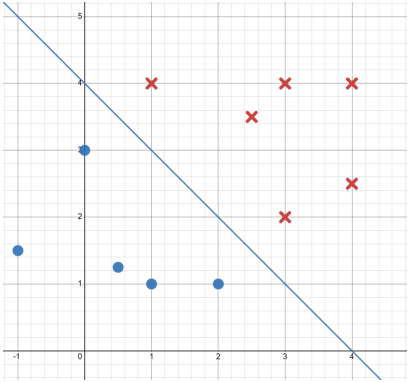
An example with a bad choice of hyperplane



What is $y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2})$ for the point $x_i = (0, 3)$?

- Blue circles: $y_i = -1$
- Red Xs: $y_i = 1$
- $-\frac{4}{\sqrt{5}} + \frac{1}{\sqrt{5}} X_1 + \frac{2}{\sqrt{5}} X_1 = 0$

Second constraint extra space



$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}) \geq M$$

- Blue circles: $y_i = -1$
- Red Xs: $y_i = 1$
- $-2\sqrt{2} + \frac{\sqrt{2}}{2}X_1 + \frac{\sqrt{2}}{2}X_1 = 0$

Mathematical Formulation

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

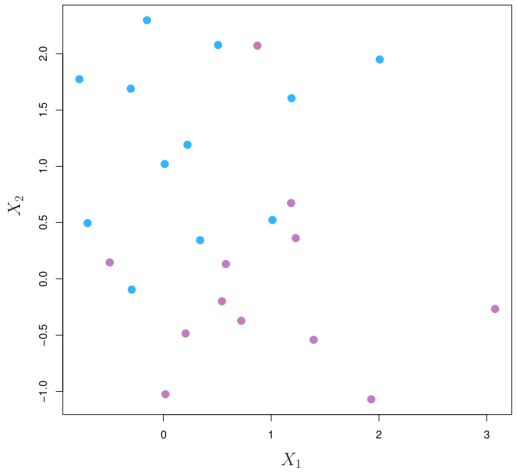
$$\text{subject to } \sum_{j=1}^{p} \beta_j^2 = 1,$$

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

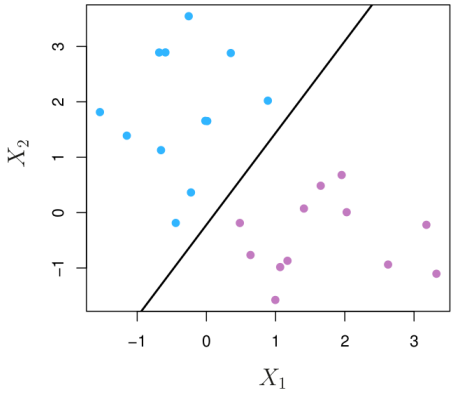
Section 2

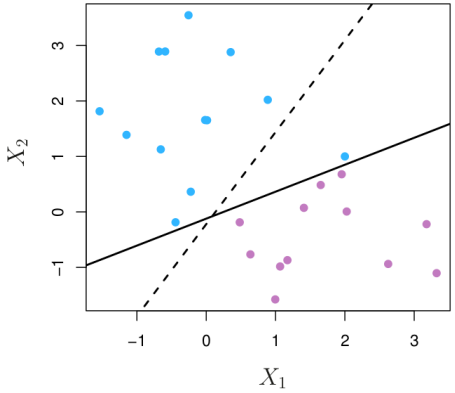
Issues with Maximal Margin Classifier

But what if....



Sensitivity to new points





Next time

21	W	3/18	Polynomial & Step Functions	7.1-7.2		
22	F	3/20	Step Functions; Basis functions; Start Splines	7.2-7.4		
23	M	3/23	Regression Splines	7.4		Q7
24	W	3/25	Decision Trees	8.1		
25	F	3/27	Random Forests	8.2.1, 8.2.2	HW #5 Due Sun 3/29	
26	M	3/30	Maximal Margin Classifier	9.1		Q8
27	W	4/1	SVC	9.2		
28	F	4/3	SVM	9.3, 9.4		
29	M	4/6	Single Layer NN	10.1		Q9
30	W	4/8	Multi Layer NN	10.2		
31	F	4/10	CNN	10.3	HW #6 Due Sun 4/12	
32	M	4/13	Unsupervised learning / clustering	12.1, 12.4		Q10
33	W	4/15	Virtual: Project Office Hours			
	F	4/17	***Review***			
	M	4/20	**Midterm #3**			
	W	4/22				
	F	4/24			**Project Due**	