Applied Text Analytics & Natural Language Processing

with Dr. Mahdi Roozbahani
& Wafa Louhichi

Support Vector Machine – Part 1

Learning Objectives

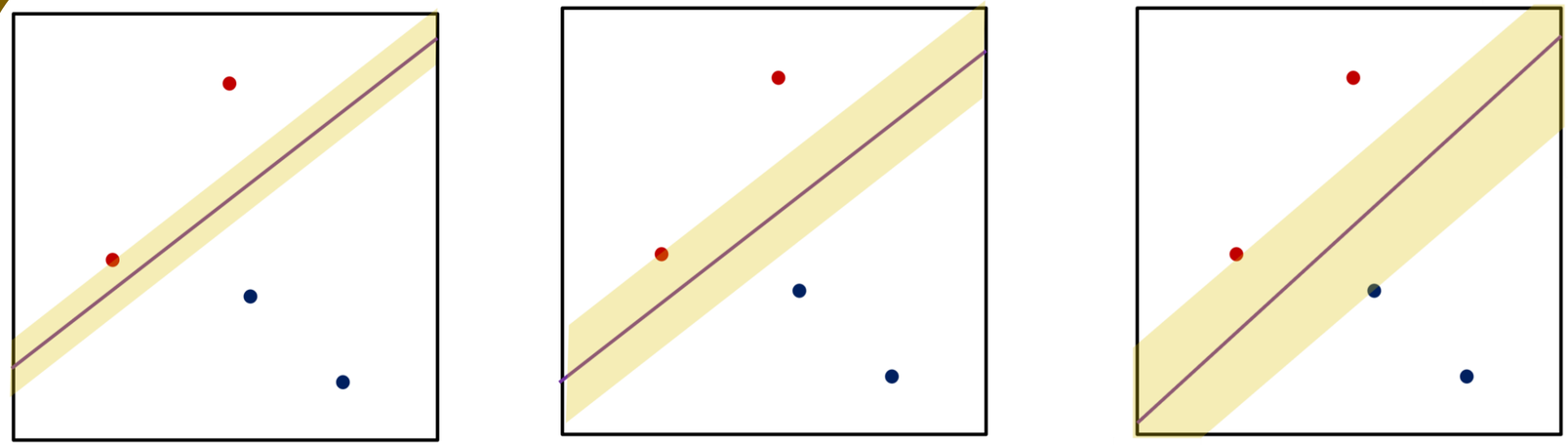
In this lesson, you will learn another linear text classifier

- Large Margin Classifier
- SVM
- Dual Form and Primal Form
- Kernel Trick

Linear Separation

We can have different separating lines

Let’s refer back to perceptron, are all these graphs a viable solution for perceptron?



Which line is the best?

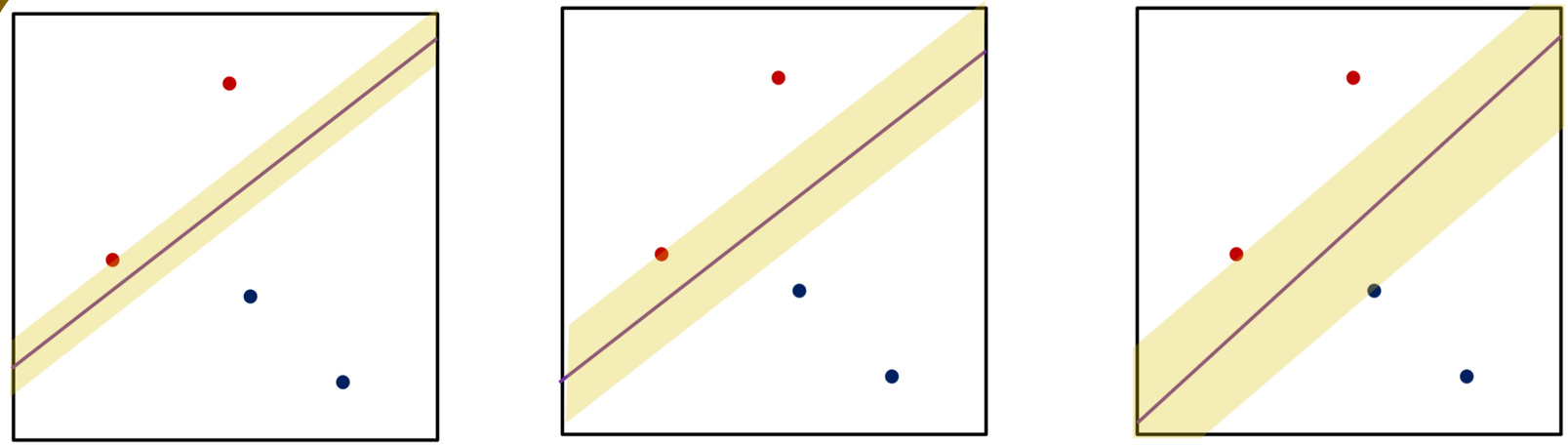
- Why is the bigger margin better?
- What $\boldsymbol{\theta}$ maximizes the margin?

All cases, error is zero and they are linear, so they are all good for generalization.

Linear Separation

We can have different separating lines

Let’s refer back to perceptron, are all these graphs a viable solution for perceptron?



Which line is the best?

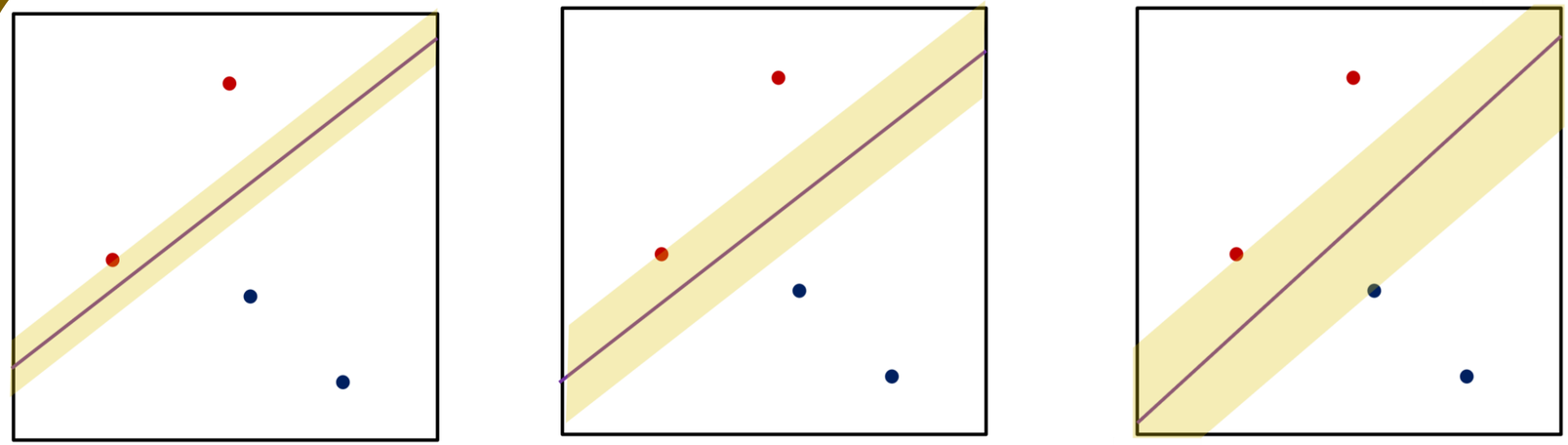
- Why is the bigger margin better?
- What θ maximizes the margin?

All cases, error is zero and they are linear, so they are all good for generalization.

Linear Separation

We can have different separating lines

Let’s refer back to perceptron, are all these graphs a viable solution for perceptron?

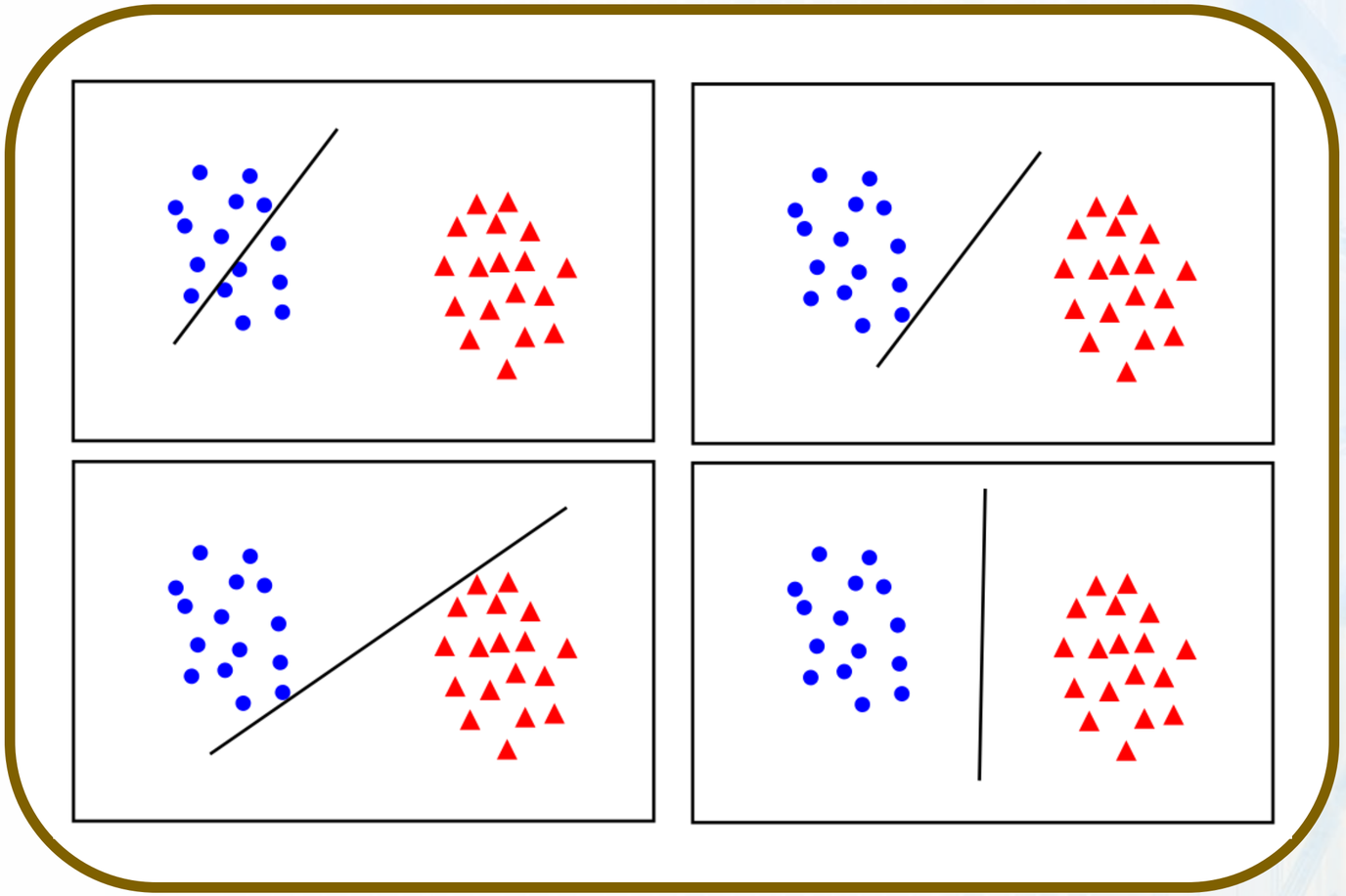


Which line is the best?

- Why is the bigger margin better?
- What $\boldsymbol{\theta}$ maximizes the margin?

All cases, error is zero and they are linear, so they are all good for generalization.

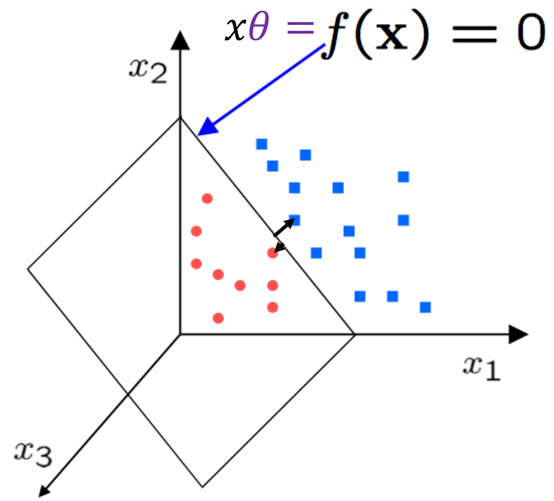
What is the Best θ?



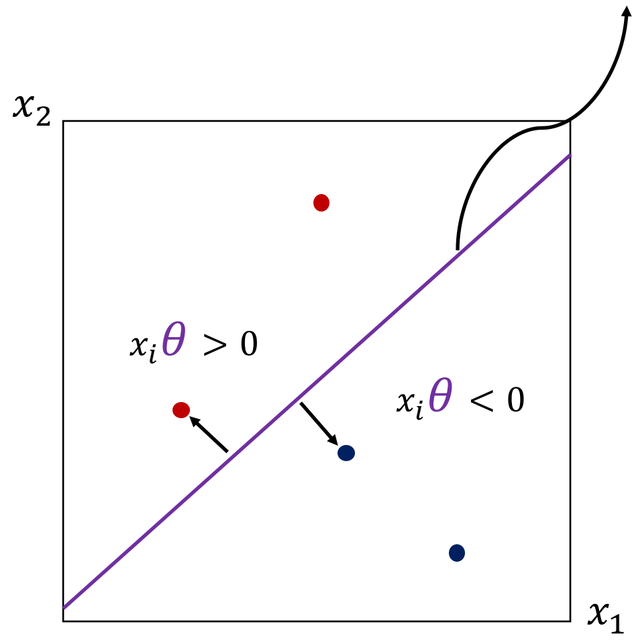
- Maximum margin solution: most table under perturbations of the inputs

Finding θ that Maximizes Margin

Solution (decision boundary) of the line: $x\theta = 0$



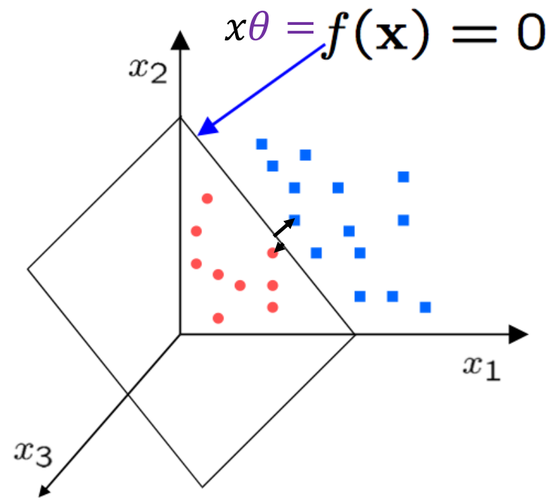
Let x_i to be the nearest data point to the line (plane):



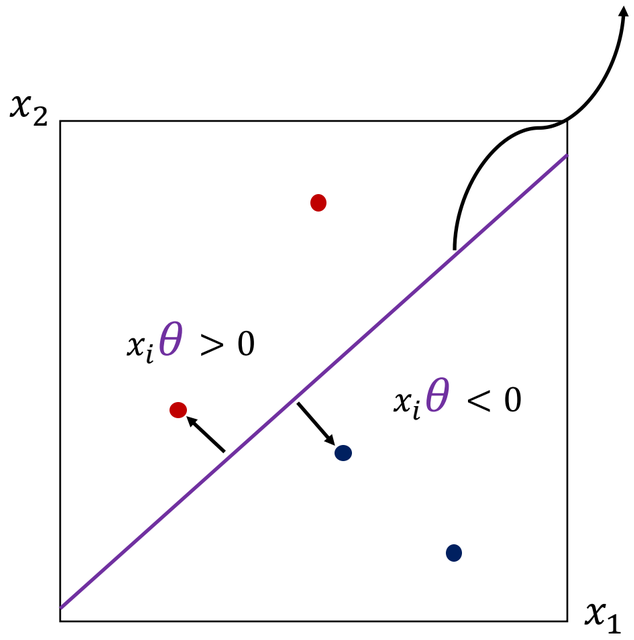
Decision boundary would be: $x\theta + b = 0$

Finding θ that Maximizes Margin

Solution (decision boundary) of the line: $x\theta = 0$



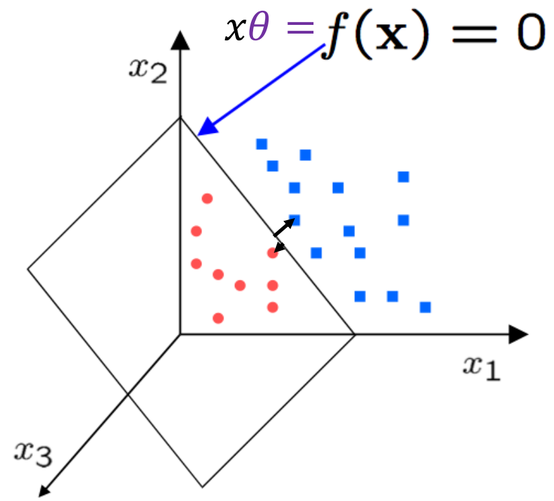
Let x_i to be the nearest data point to the line (plane):



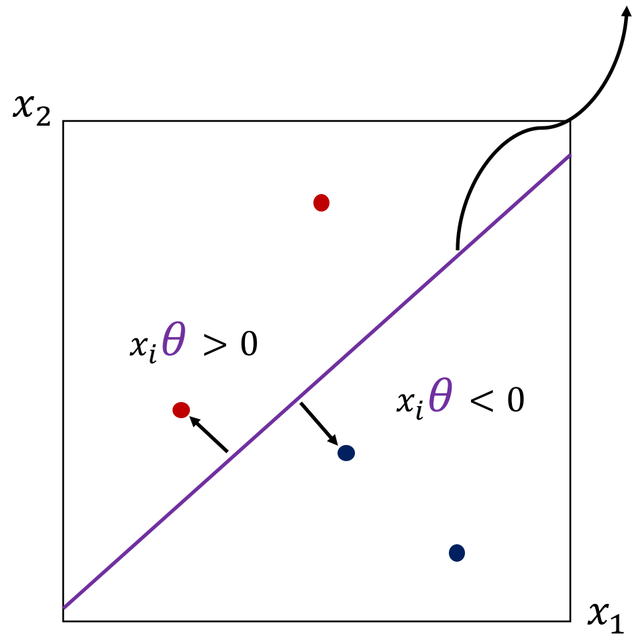
Decision boundary would be: $x\theta + b = 0$

Finding θ that Maximizes Margin

Solution (decision boundary) of the line: $x\theta = 0$



Let x_i to be the nearest data point to the line (plane):



Decision boundary would be: $x\theta + b = 0$

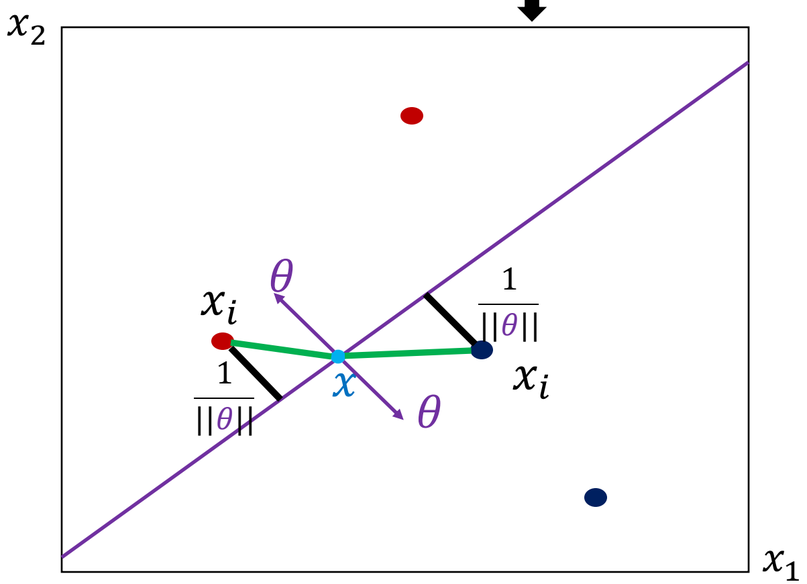
What is the Length of My Large Margin?

$$\text{distance} = \frac{1}{||\theta||}|(x_i\theta - x\theta)|$$

$$= \frac{1}{||\theta||}|(\underbrace{x_i\theta + b}_{\text{My constraint}} \underbrace{- x\theta - b}_{\text{A point on the decision line}})|$$

$$|x_i\theta + b| = 1 \qquad x\theta + b = 0$$

$$= \frac{2}{||\theta||}$$ The total margin



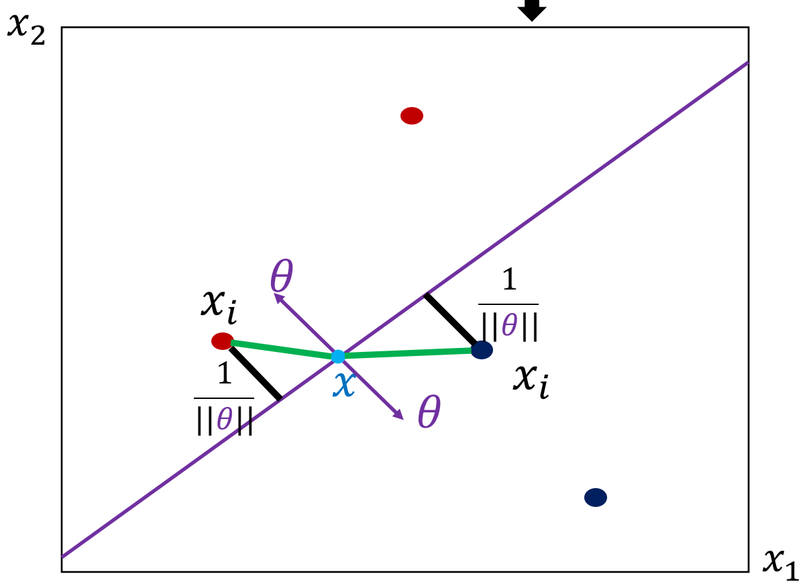
What is the Length of My Large Margin?

$$\text{distance} = \frac{1}{||\theta||}|(x_i\theta - x\theta)|$$

$$= \frac{1}{||\theta||}|(\underbrace{x_i\theta + b}_{\text{My constraint}} \underbrace{- x\theta - b}_{\text{A point on the decision line}})|$$

$$|x_i\theta + b| = 1 \qquad x\theta + b = 0$$

$$= \frac{2}{||\theta||}$$ The total margin



What is the Length of My Large Margin?

$$\text{distance} = \frac{1}{||\theta||}|(x_i\theta - x\theta)|$$

$$= \frac{1}{||\theta||}|(\underbrace{x_i\theta + b}_{\text{My constraint}} \underbrace{- x\theta - b}_{\text{A point on the decision line}})|$$

$$|x_i\theta + b| = 1 \qquad x\theta + b = 0$$

$$= \frac{2}{||\theta||}$$ The total margin