Kernel methods I
Basis expansion

Sanjoy Dasgupta

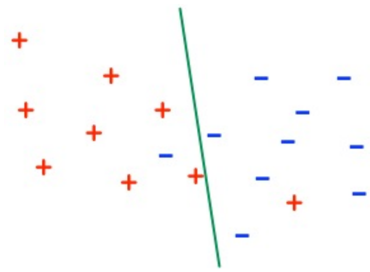
University of California, San Diego

Topics we'll cover

1. Two deviations from linear separability
2. Learning quadratic boundaries using basis expansion

Deviations from linear separability

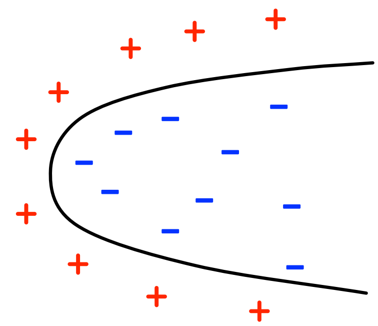
Noise



Find a separator that minimizes a convex loss function related to the number of mistakes.

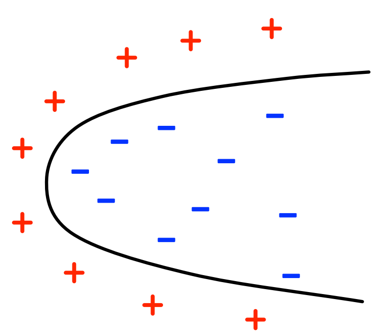
e.g. SVM, logistic regression.

Systematic deviation



What to do with this?

Adding new features



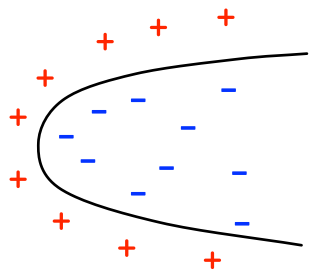
Actual boundary is something like $x_1 = x_2^2 + 5$.

- This is quadratic in $x = (x_1, x_2)$
- But it is linear in $\Phi(x) = (x_1, x_2, x_1^2, x_2^2, x_1 x_2)$

Basis expansion: embed data in a higher-dimensional feature space. Then we can use a linear classifier!

Basis expansion for quadratic boundaries

How to deal with a **quadratic** boundary?



Idea: augment the regular features $x = (x_1, x_2, \ldots, x_d)$ with

$$x_1^2,\ x_2^2,\ \ldots,\ x_d^2$$
$$x_1x_2,\ x_1x_3,\ \ldots,\ x_{d-1}x_d$$

Enhanced data vectors of the form:

$$\Phi(x) = (x_1, \ldots, x_d, x_1^2, \ldots, x_d^2, x_1x_2, \ldots, x_{d-1}x_d)$$

Quick question

Suppose $x = (x_1, x_2, x_3)$. What is the dimension of $\Phi(x)$?

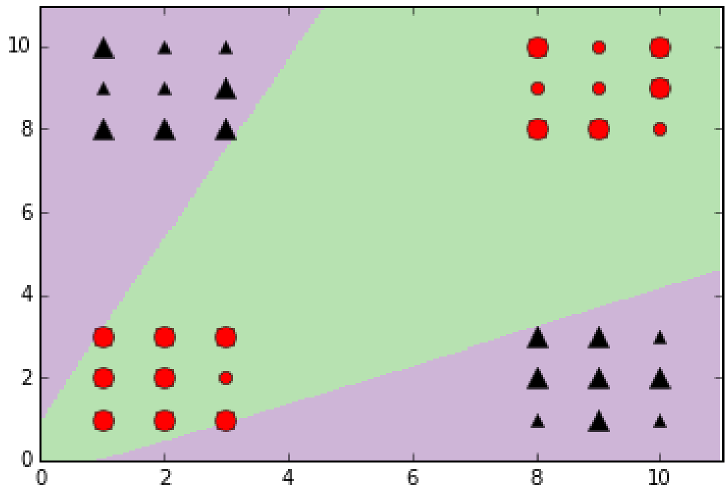
Suppose $x = (x_1, \ldots, x_d)$. What is the dimension of $\Phi(x)$?

Perceptron revisited

Learning in the higher-dimensional feature space:

- $w = 0$ and $b = 0$
- while some $y(w \cdot \Phi(x) + b) \leq 0$:
  - $w = w + y\,\Phi(x)$
  - $b = b + y$

Perceptron with basis expansion: examples



Perceptron with basis expansion: examples