Lecture 13E. An Extra Bit on Linear SVM

COMP90051 Statistical Machine Learning

Semester 2, 2015
Lecturer: Andrey Kan

THE UNIVERSITY OF MELBOURNE

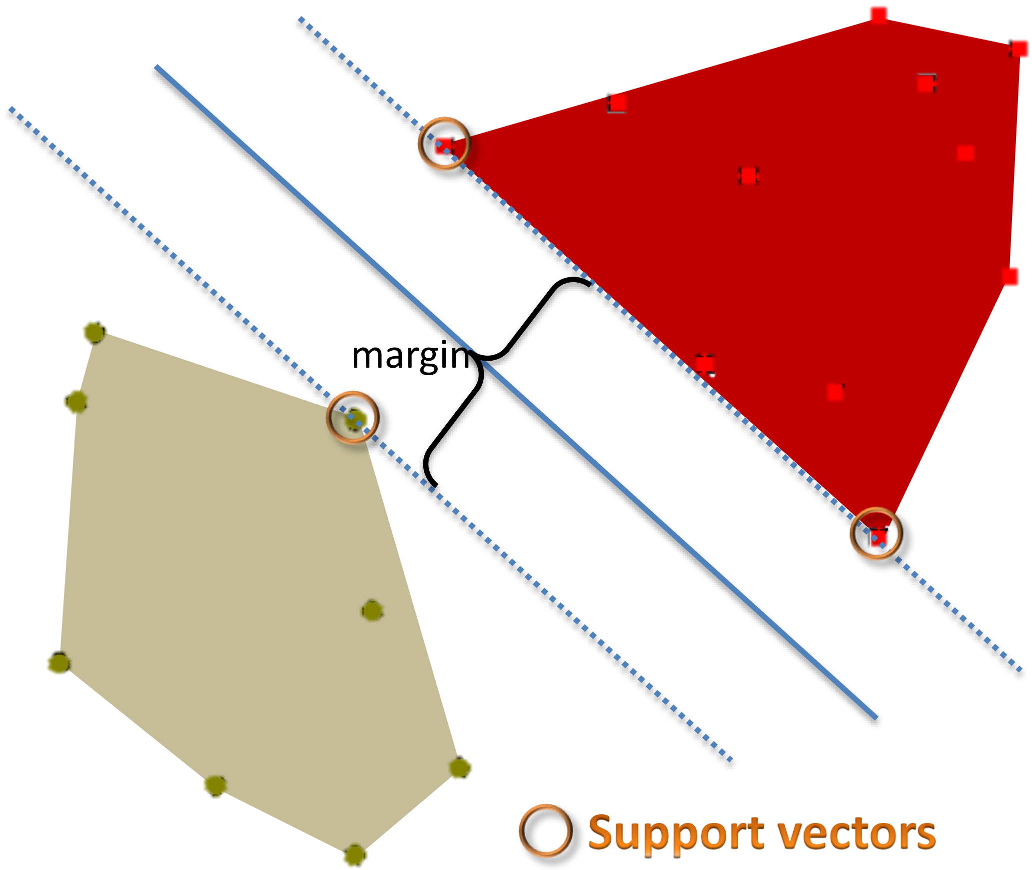
Do Ask Questions!

- If something doesn’t make sense
- If I’m speaking too fast or too slow
- If you have any other concerns

- In class or just after the lecture
- During the office hour
- Via email (to me or Ben)
- Via the LMS forum
- Anonymously via the LMS forum

Support Vector Machines: Intuition

- What linear classifier to choose?

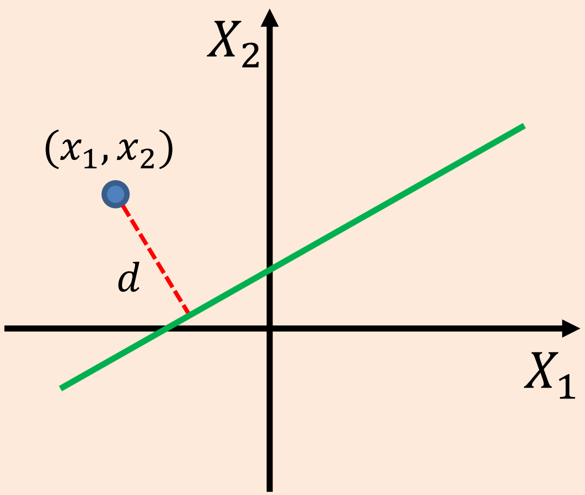


Better!

Maximise margin leaving plenty of room for noise
→ train, test error

(very brief) Revision of geometry

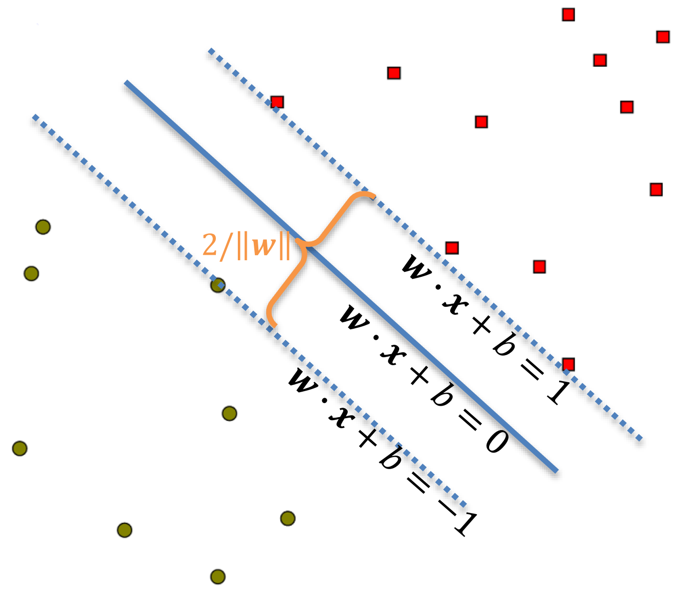
- Line = set of points (x_1, x_2) that satisfy $w_1 x_1 + w_2 x_2 + b = 0$
- Plane ➔ $w_1 x_1 + w_2 x_2 + w_3 x_3 + b = 0$
- Hyperplane ➔ $\sum_{i=1}^{n} w_i x_i + b = 0$



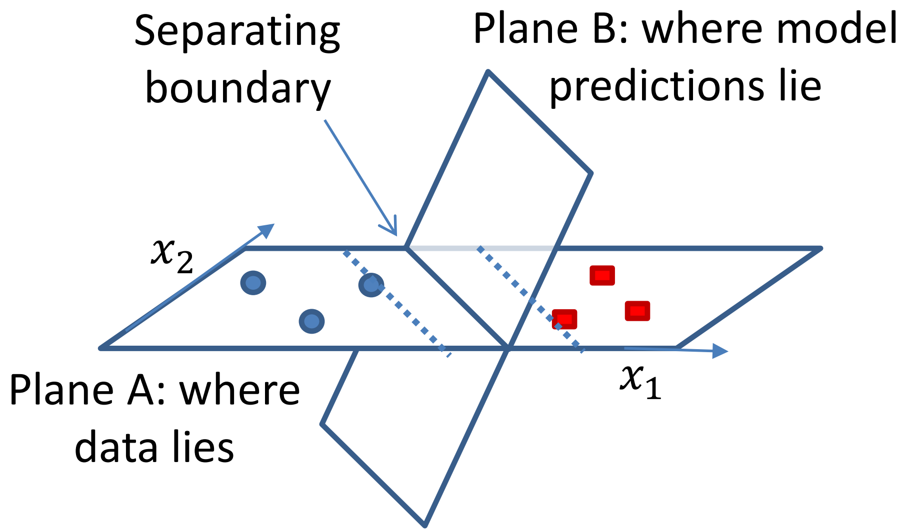
- <u>Note</u>: we can multiply w_i, b by a constant without changing the (hyper-)plane
- Distance from point $\boldsymbol{x}$ to (hyper-)plane is $d = \frac{|\boldsymbol{w}^T \boldsymbol{x} + b|}{\|\boldsymbol{w}\|}$
- We can rescale so that $\boldsymbol{w}^T \boldsymbol{x} + b = 1$, and get $d = \frac{1}{\|\boldsymbol{w}\|}$

Linear SVM parameters

- Goal: max margin hyperplane
- Hyperplane parameters
  - Normal vector $\boldsymbol{w}$
  - Offset b
- Scale parameters so that
  - Negative data: $\boldsymbol{w} \cdot \boldsymbol{x} + b \leq -1$
  - Positive data: $\boldsymbol{w} \cdot \boldsymbol{x} + b \geq +1$
- Then margin is $2/\|\boldsymbol{w}\|$

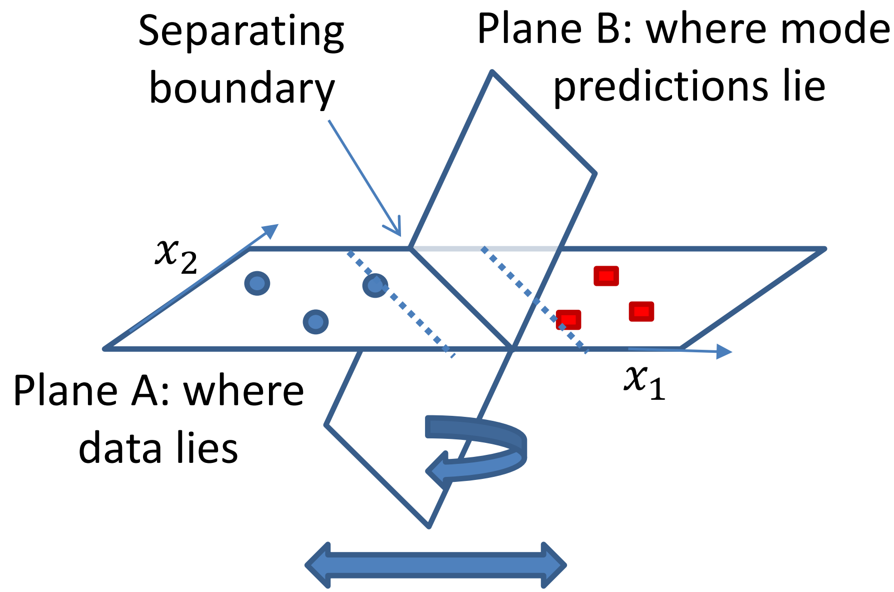


- Consider 2D data (with two features x_1, x_2)



Dashed lines show the margin

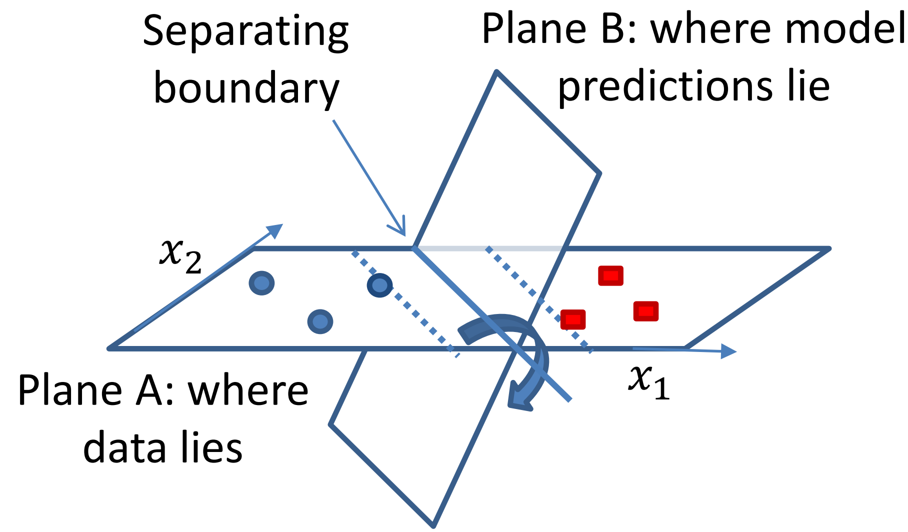
- Training a linear SVM essentially means moving/rotating plane B so that separating boundary changes



- However, we can also rotate Plane B along the separating boundary
- In this case, the boundary does NOT change
  - Same classifier!

➔ The same boundary can be expressed by many parameter combinations

Each corresponding to a particular rotation of Plane B

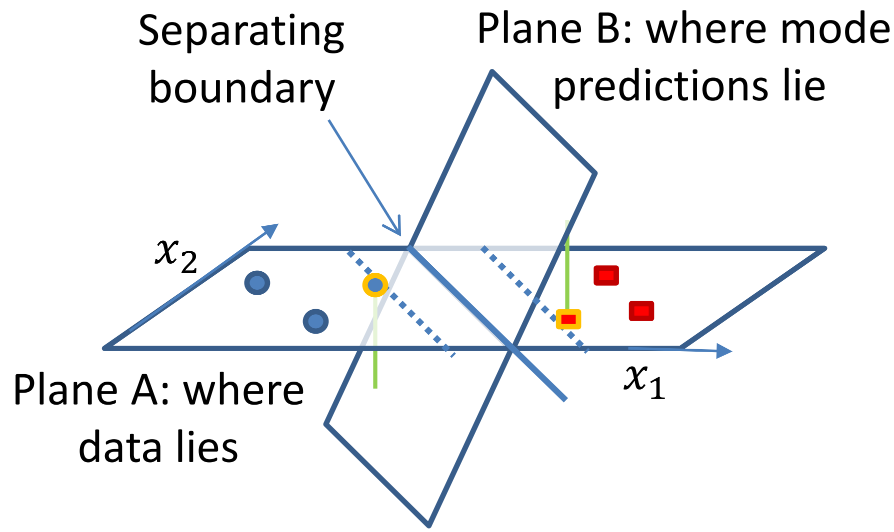


- However, we can also rotate Plane B along the separating boundary
- In this case, the boundary does NOT change
  - Same classifier!

➔ The same boundary can be expressed by many parameter combinations

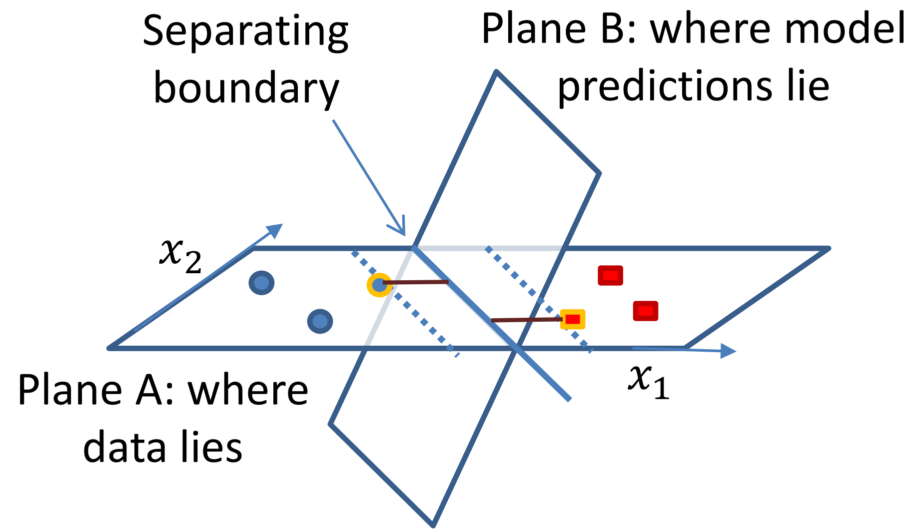
Each corresponding to a particular rotation of Plane B

- For convenience we choose a particular Plane B angle so that the distance from a support vector to Plane B equals to 1



Dashed lines show the margin
Support vectors are marked with golden outline
Green distances = 1

- For convenience we choose a particular Plane B angle so that the distance from a support vector to Plane B equals to 1
- In this case, the width of the margin is $\frac{2}{\|\boldsymbol{w}\|}$



Dashed lines show the margin
Support vectors are marked with golden outline
Brown distances = $\frac{1}{\|\boldsymbol{w}\|}$

Linear SVM parameters

- Goal: max margin hyperplane
- Hyperplane parameters
  - Normal vector $\boldsymbol{w}$
  - Offset b
- Scale parameters so that
  - Negative data: $\boldsymbol{w} \cdot \boldsymbol{x} + b \leq -1$
  - Positive data: $\boldsymbol{w} \cdot \boldsymbol{x} + b \geq +1$
- Then margin is $2/\|\boldsymbol{w}\|$