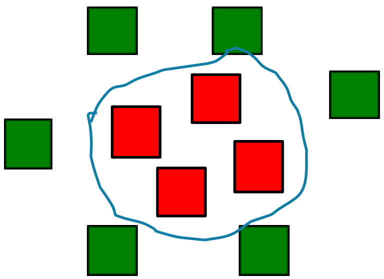
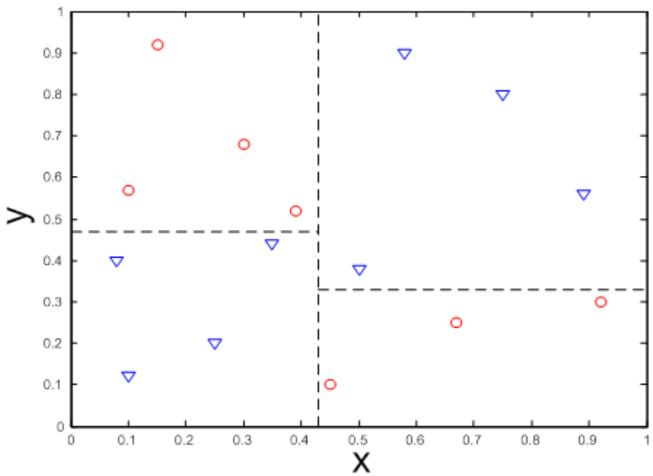
MAX-MARGIN SVMS

SYRACUSE UNIVERSITY
School of Information Studies

THE SHAPE OF DECISION BOUNDARY

Some data are not linearly separable.

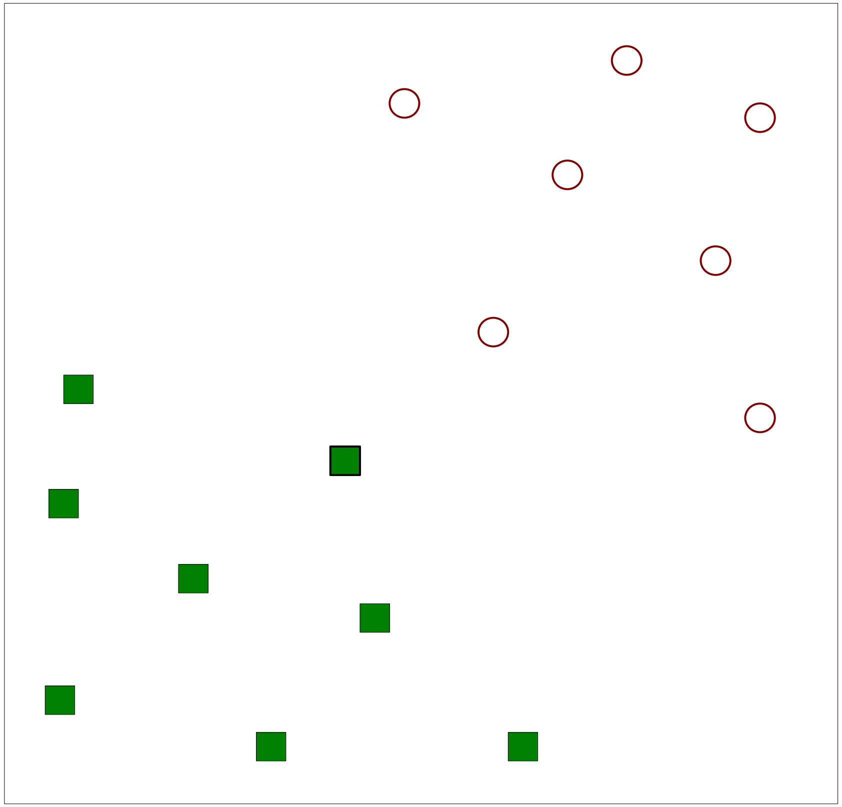




Support vector machine (SVM): An algorithm that can solve both linearly separable and inseparable problems

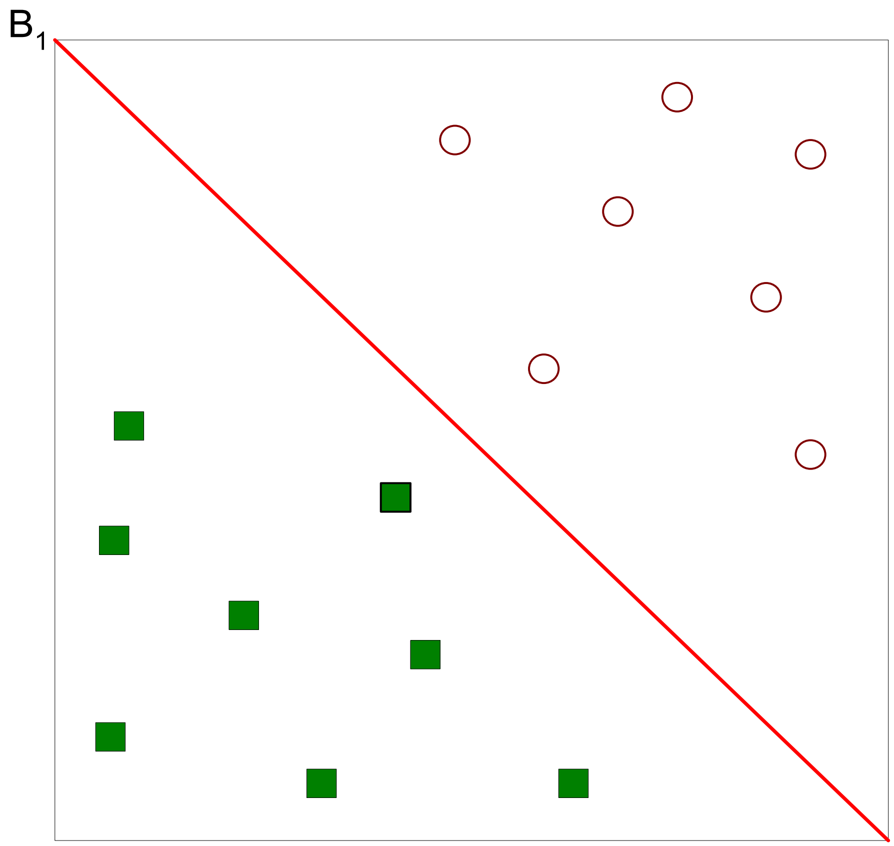
SUPPORT VECTOR MACHINES

Find a linear hyperplane (decision boundary) that can separate the data.



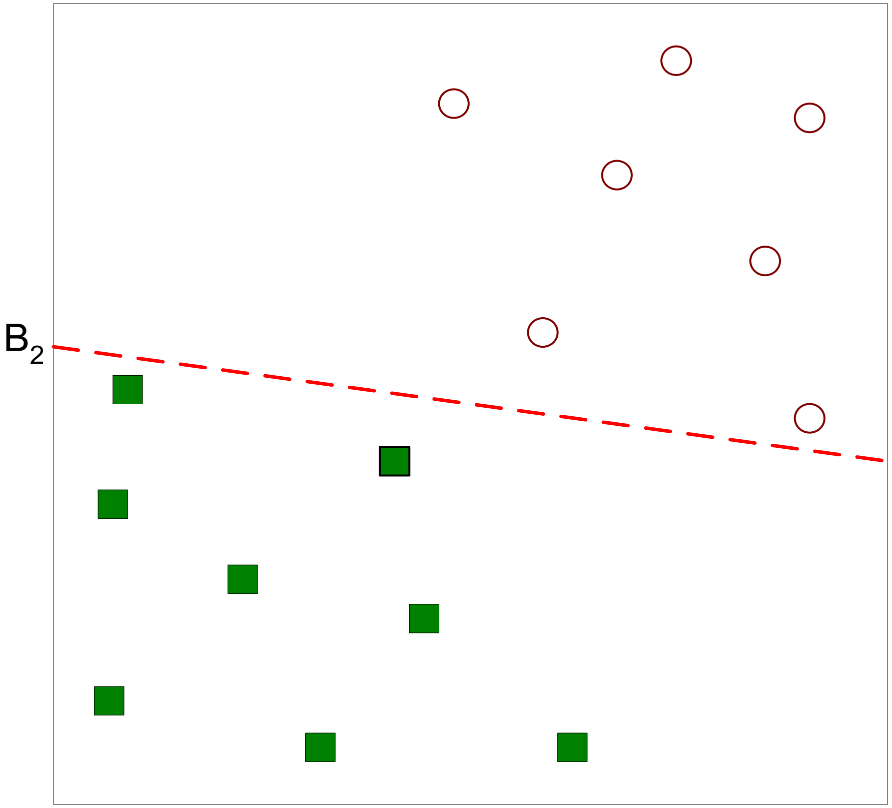
SUPPORT VECTOR MACHINES

One possible solution:



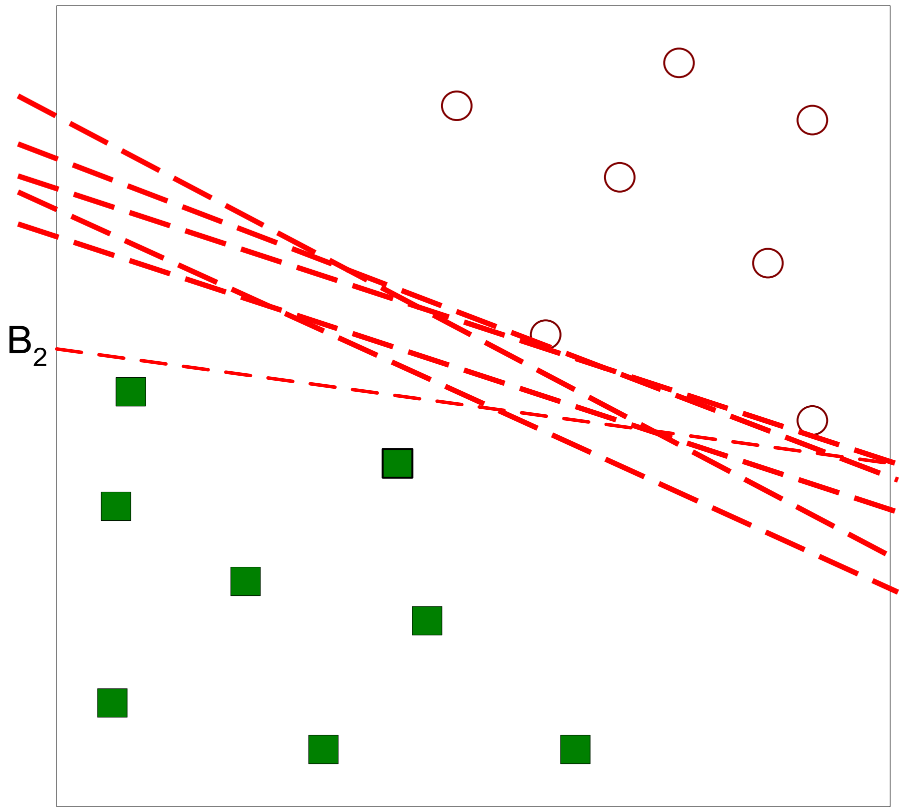
SUPPORT VECTOR MACHINES

Another possible solution:



SUPPORT VECTOR MACHINES

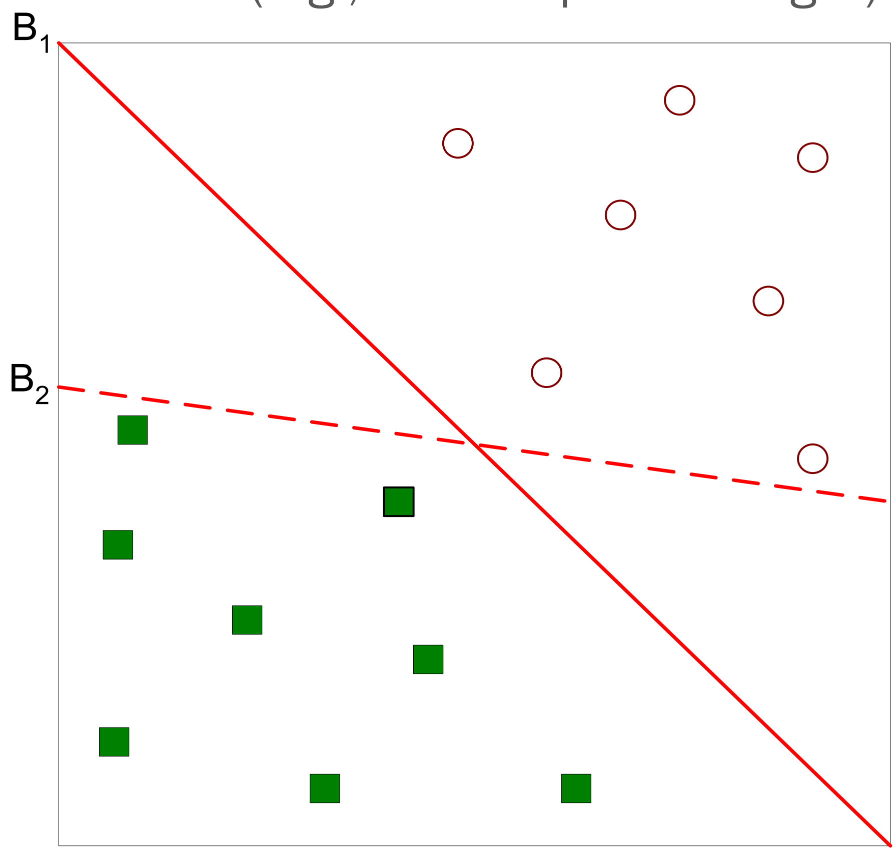
Numerous possible solutions:



SUPPORT VECTOR MACHINES

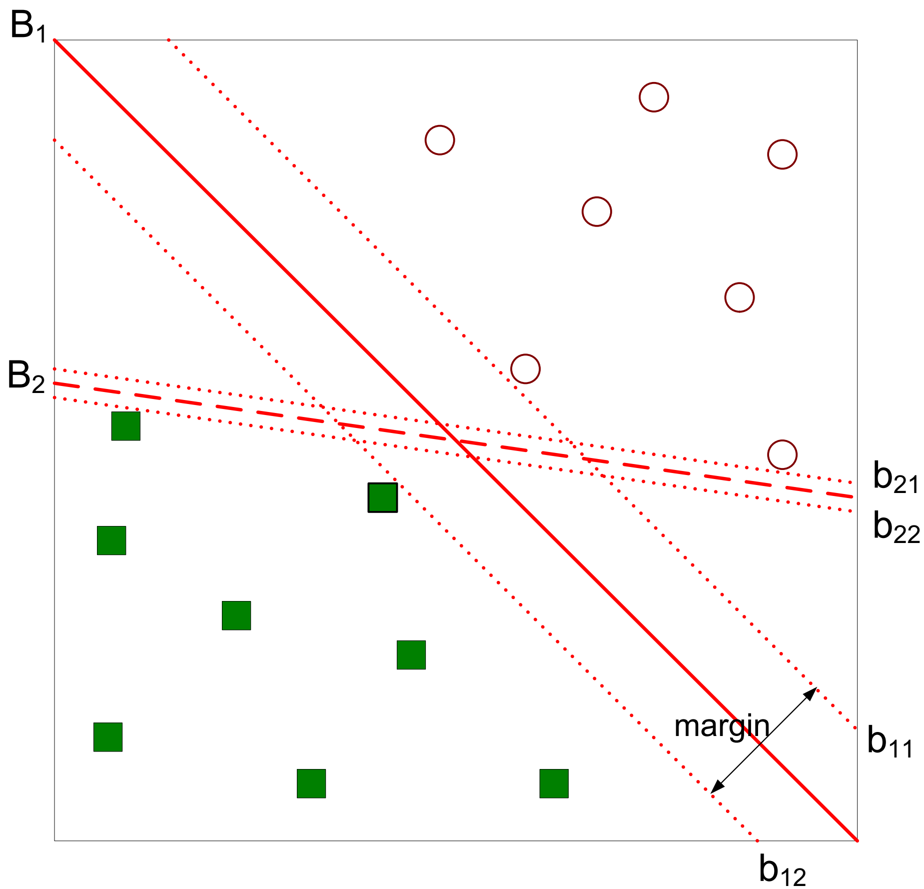
Which one is better: B1 or B2?

How do you define better? (E.g., “least square fitting”?)

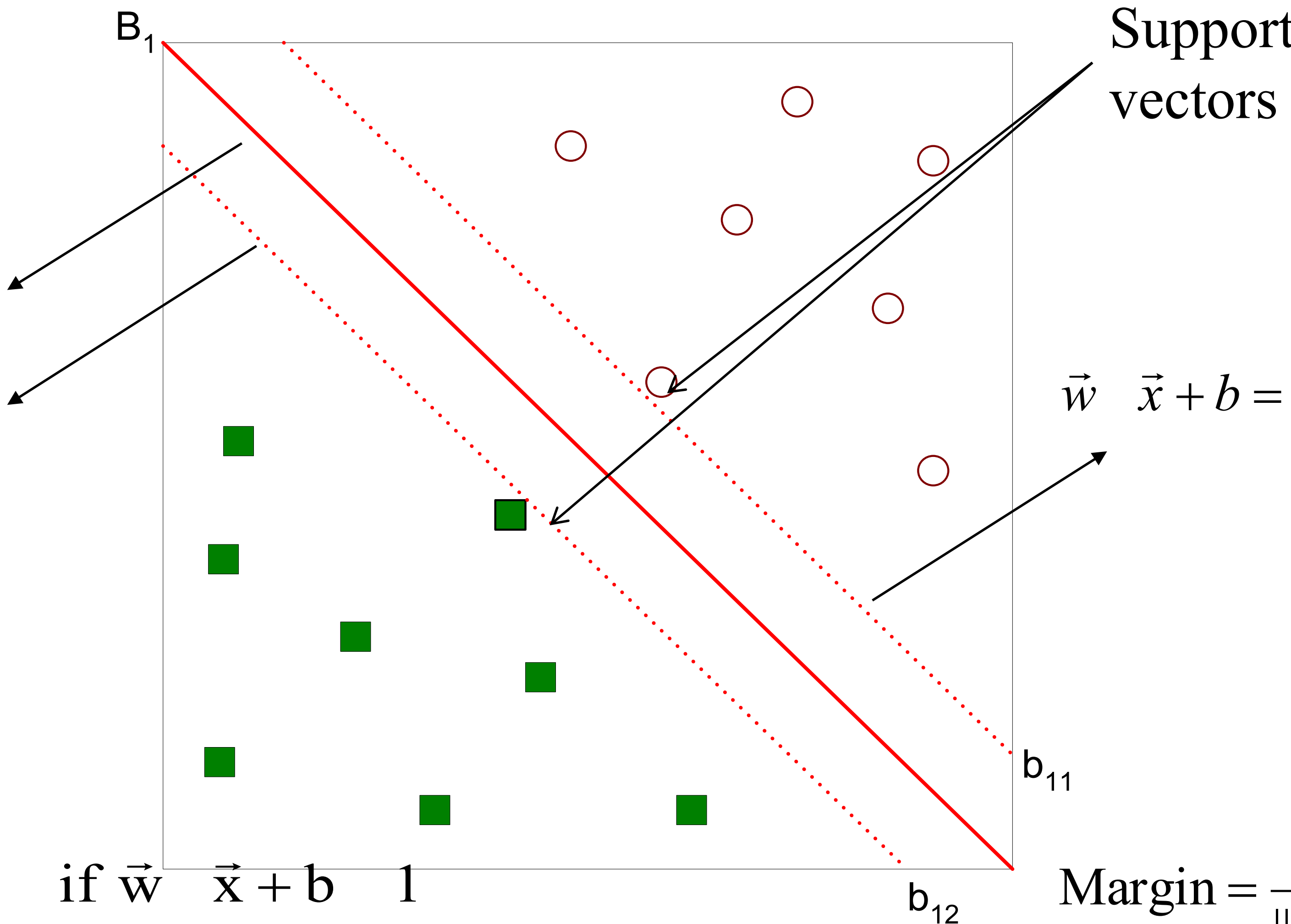


SUPPORT VECTOR MACHINES

Find a hyperplane that maximizes the margin => B1 is better than B2.



SUPPORT VECTOR MACHINES



$$\vec{w} \quad \vec{x} + b = 0$$

$$\vec{w} \quad \vec{x} + b = \quad 1$$

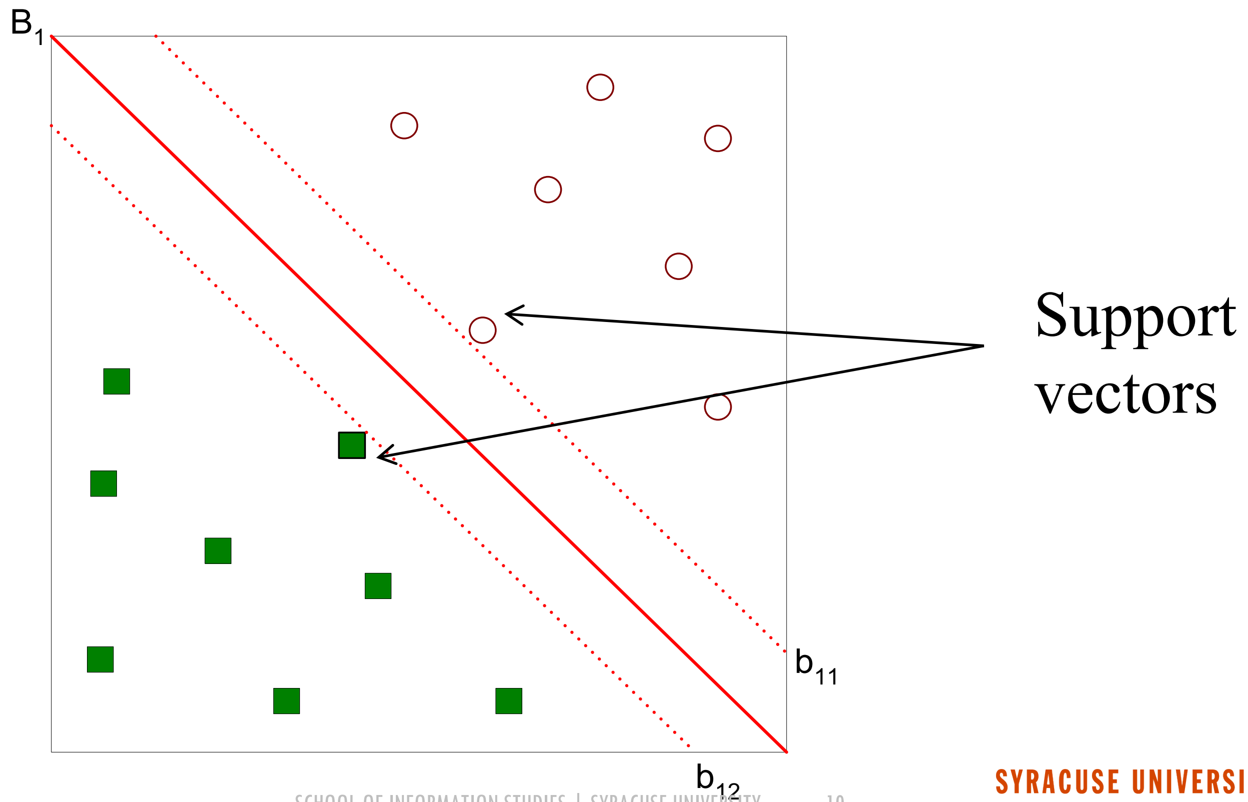
$$\vec{w} \quad \vec{x} + b = +1$$

$$f(\vec{x}) = \begin{array}{ll} 1 & \text{if } \vec{\mathrm{w}} \quad \vec{\mathrm{x}} + \mathrm{b} \quad 1 \\ 1 & \text{if } \vec{\mathrm{w}} \quad \vec{\mathrm{x}} + \mathrm{b} \quad 1 \end{array}$$

$$\text{Margin} = \frac{2}{\| \vec{w} \|^2}$$

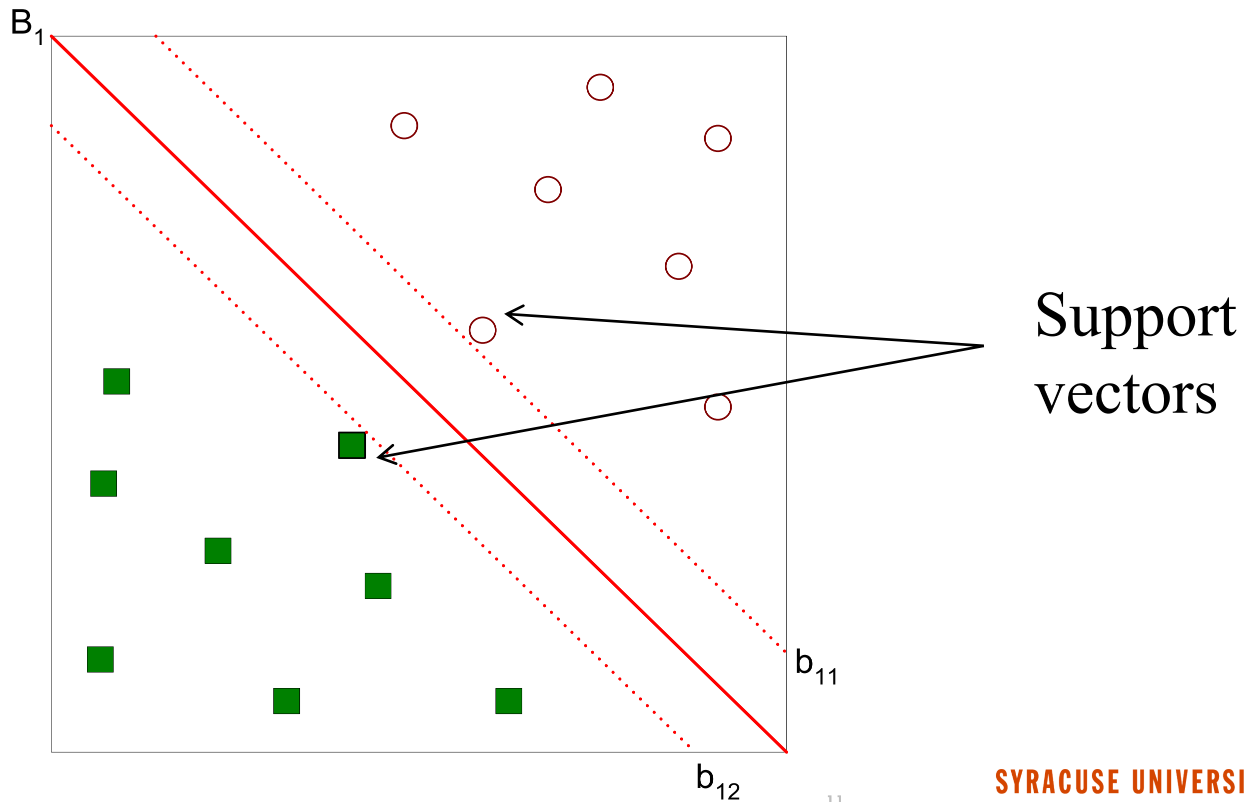
SUPPORT VECTORS

Support vectors are the training examples ("vectors") that are located on the margins.



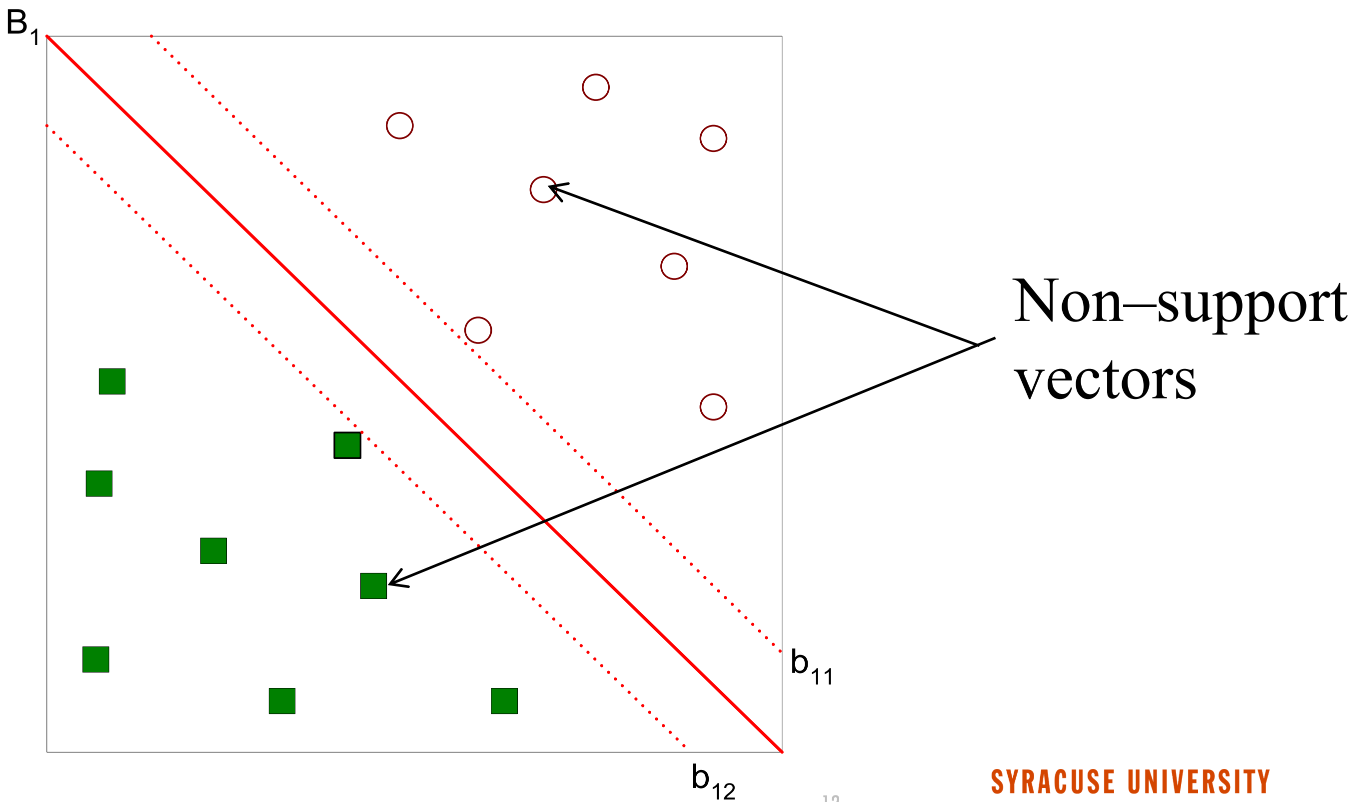
DECISION BOUNDARY

Only support vectors determine the decision boundary.



NON–SUPPORT VECTORS

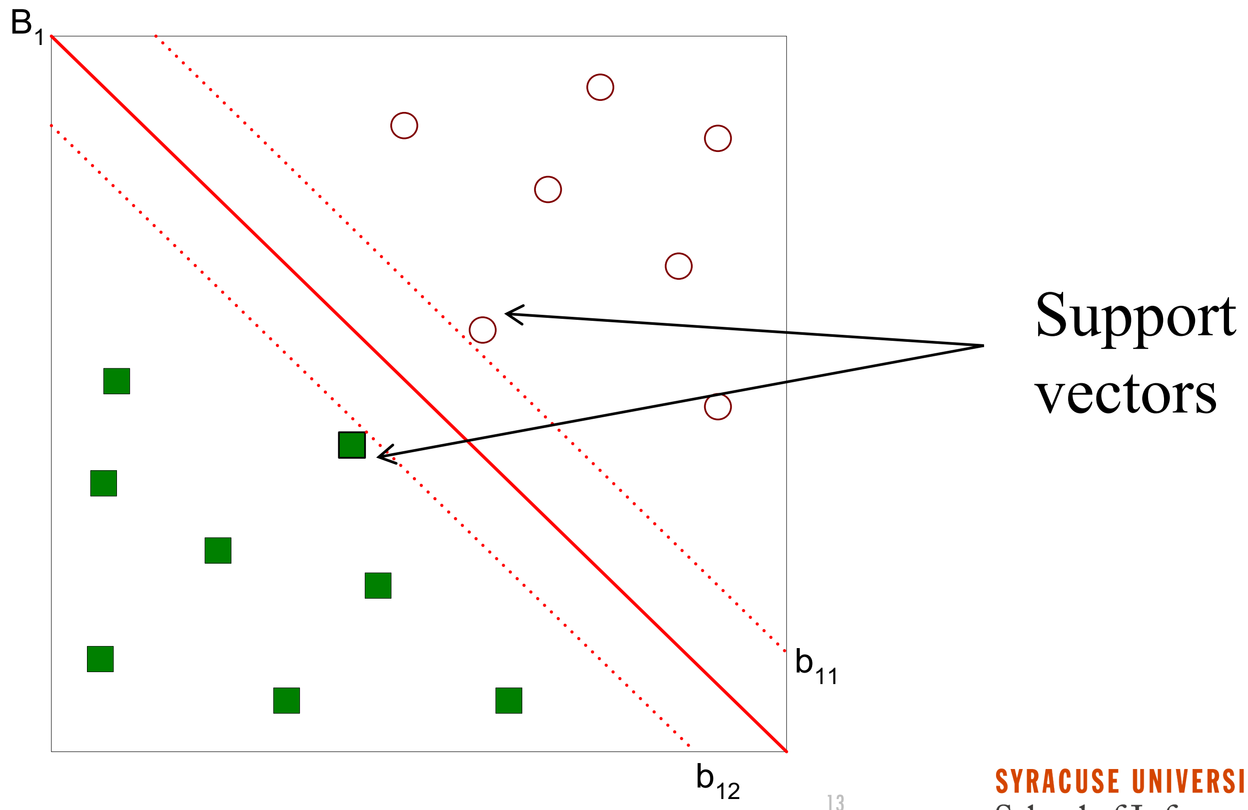
Training examples that are not support vectors do not participate in prediction.



SYRACUSE UNIVERSITY
School of Information Studies

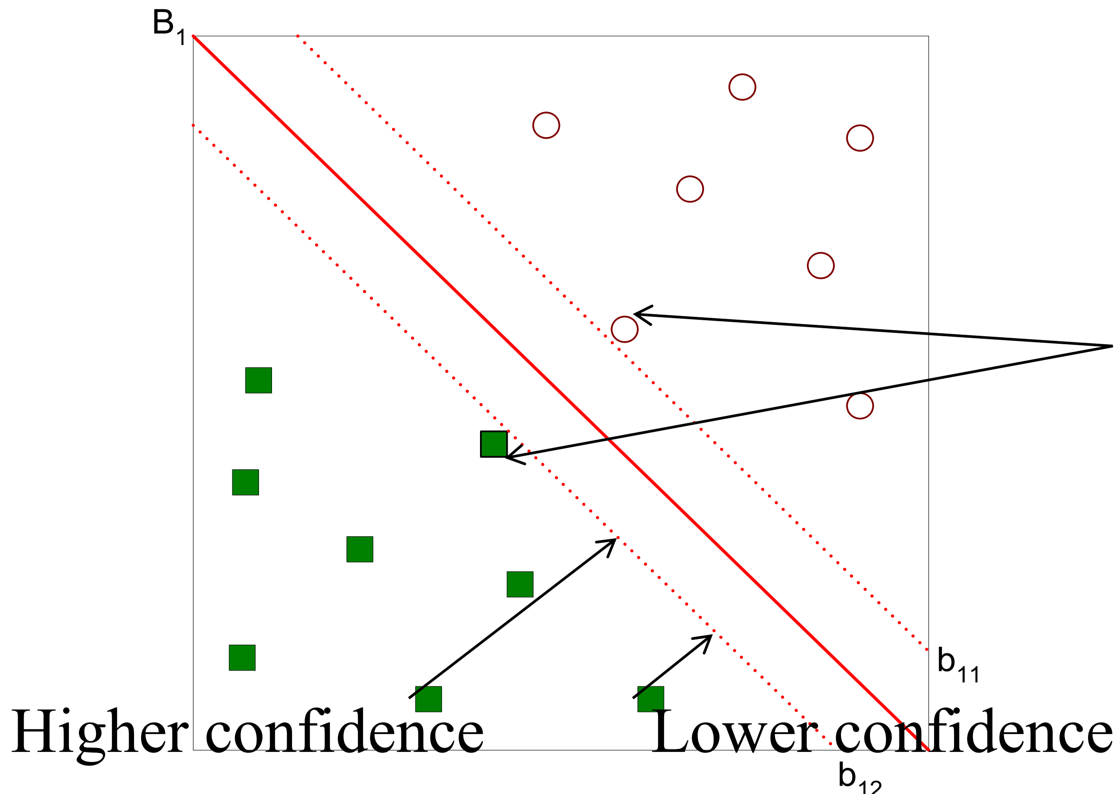
MODEL COMPLEXITY

The number of support vectors is an indicator of the complexity of the trained SVM model.



PREDICTION CONFIDENCE

The distance between the example and the decision boundary is an indicator of prediction confidence: The farther the better.



PREDICTION CONFIDENCE

SVM's prediction result can be sorted by confidence and thus is suitable for semi-supervised learning and active learning.

Variant SVMs algorithm can be used for regression.