

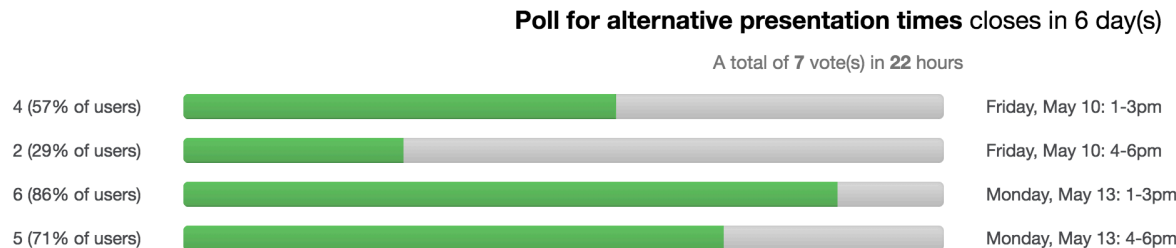
CS 66: Machine Learning

Prof. Sara Mathieson

Spring 2019



- CR/NC deadline TODAY
 - If you're considering this, we can discuss in office hours today (or before/after)
- Poll for Alternative presentation times



- Tentative schedule
 - Lab 6 due April 5
 - Lab 7 due April 15
 - Project Proposal due April 19
 - Midterm 2 on April 24

Outline for March 29

- Lab 3 followups
- SVM optimization
 - Coordinate ascent
 - Incremental algorithm
- Next week: Start Neural Networks
 - Reading posted

Outline for March 29

- Lab 3 followups
- SVM optimization
 - Coordinate ascent
 - Incremental algorithm
- Next week: Start Neural Networks
 - Reading posted

Lab 3

(solutions not posted online)

Lab 3, Part (a)

- Always good to start by identifying the goal and the variable(s) to predict. Classification or Regression?
- Think about train and test holistically, not separately. Normally will not have them separated or be able to visualize them.

Lab 3, Part (e)

- As α decreases, we generally see the SGD solution become closer to the analytic solution
- But if α becomes *really* small, we see it get further away.... Why?

Lab 3, Part (g)

- RMSE has several advantages over cost, even though it is very closely related! Why?

Outline for March 29

- Lab 3 followups
- SVM optimization
 - Coordinate ascent
 - Incremental algorithm
- Next week: Start Neural Networks
 - Reading posted

Coordinate Ascent

Loop until convergence: {

For $i = 1, \dots, n$ {

$$\alpha_i \leftarrow \arg \max_{\hat{\alpha}_i} W(\alpha_1, \dots, \alpha_{i-1}, \hat{\alpha}_i, \alpha_{i+1}, \dots, \alpha_n)$$

}

}

Coordinate Ascent

```
Loop until convergence: {  
  
    For  $i = 1, \dots, n$  {  
  
         $\alpha_i \leftarrow \arg \max_{\hat{\alpha}_i} W(\alpha_1, \dots, \alpha_{i-1}, \hat{\alpha}_i, \alpha_{i+1}, \dots, \alpha_n)$   
  
    }  
  
}
```

- In practice, need to choose pairs of alphas since we have the constraint:

$$\sum_{i=1}^n y_i \alpha_i = 0$$

More details: see
Andrew Ng notes

Meta-optimization process

- Incremental SVM optimization algorithm

Meta-optimization process

- Incremental SVM optimization algorithm
- Choose a subset S of examples and run coordinate ascent fully

Meta-optimization process

- Incremental SVM optimization algorithm
- Choose a subset S of examples and run coordinate ascent fully
- Identify which alpha values are 0 \Rightarrow these cannot be support vectors in final solution!

Meta-optimization process

- Incremental SVM optimization algorithm
- Choose a subset S of examples and run coordinate ascent fully
- Identify which alpha values are 0 \Rightarrow these cannot be support vectors in final solution!
- Discard these points and add new ones; repeat

Incremental SVM Optimization Algorithm

* K = subset size

* S = subset

$$|S| = K$$

* randomize order of examples

* initially $S = \{x_1, x_2, \dots, x_K\}$

Repeat until no more examples:

- run coordinate ascent with train data S

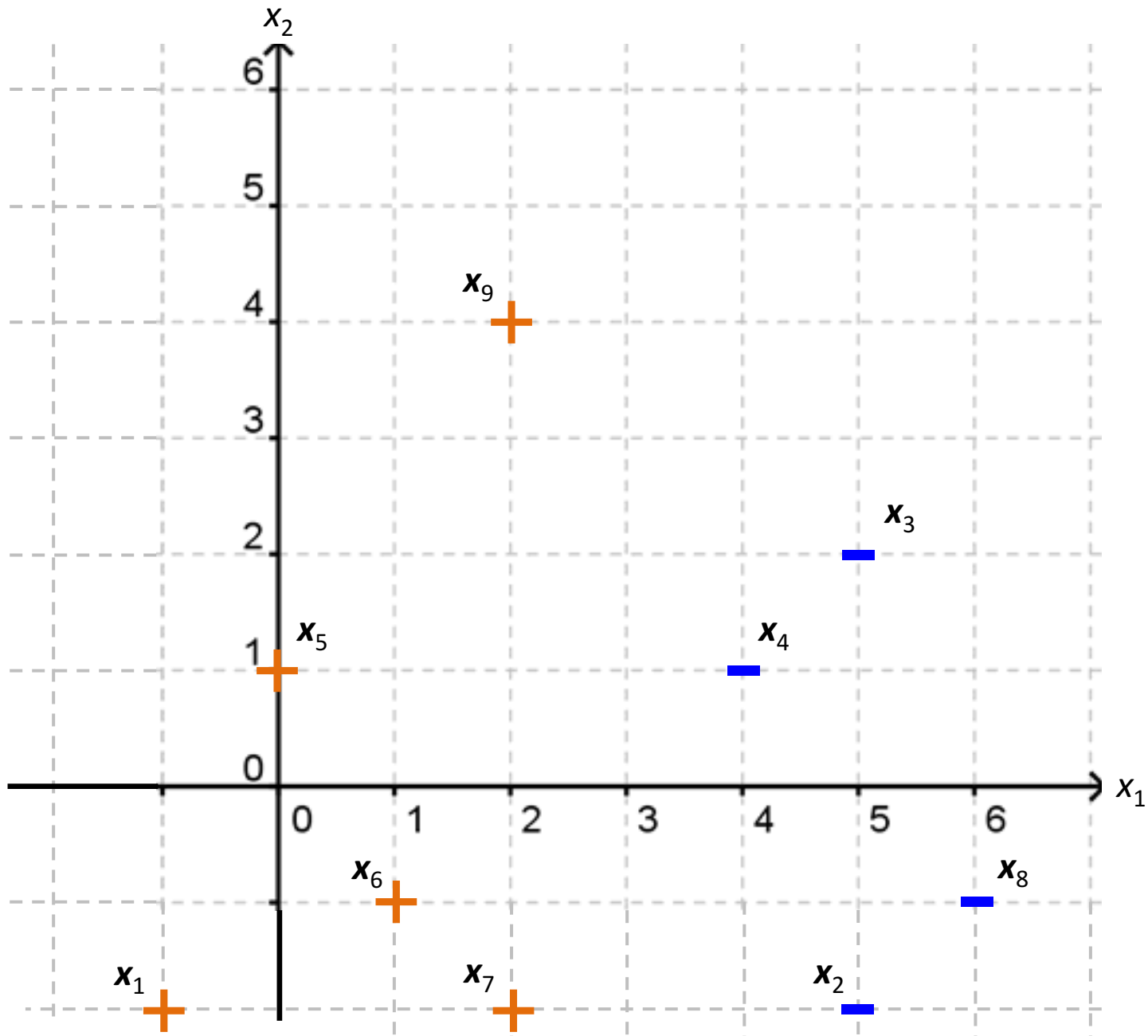
long runtime!

- \Rightarrow α 's on all these examples + hyperplane

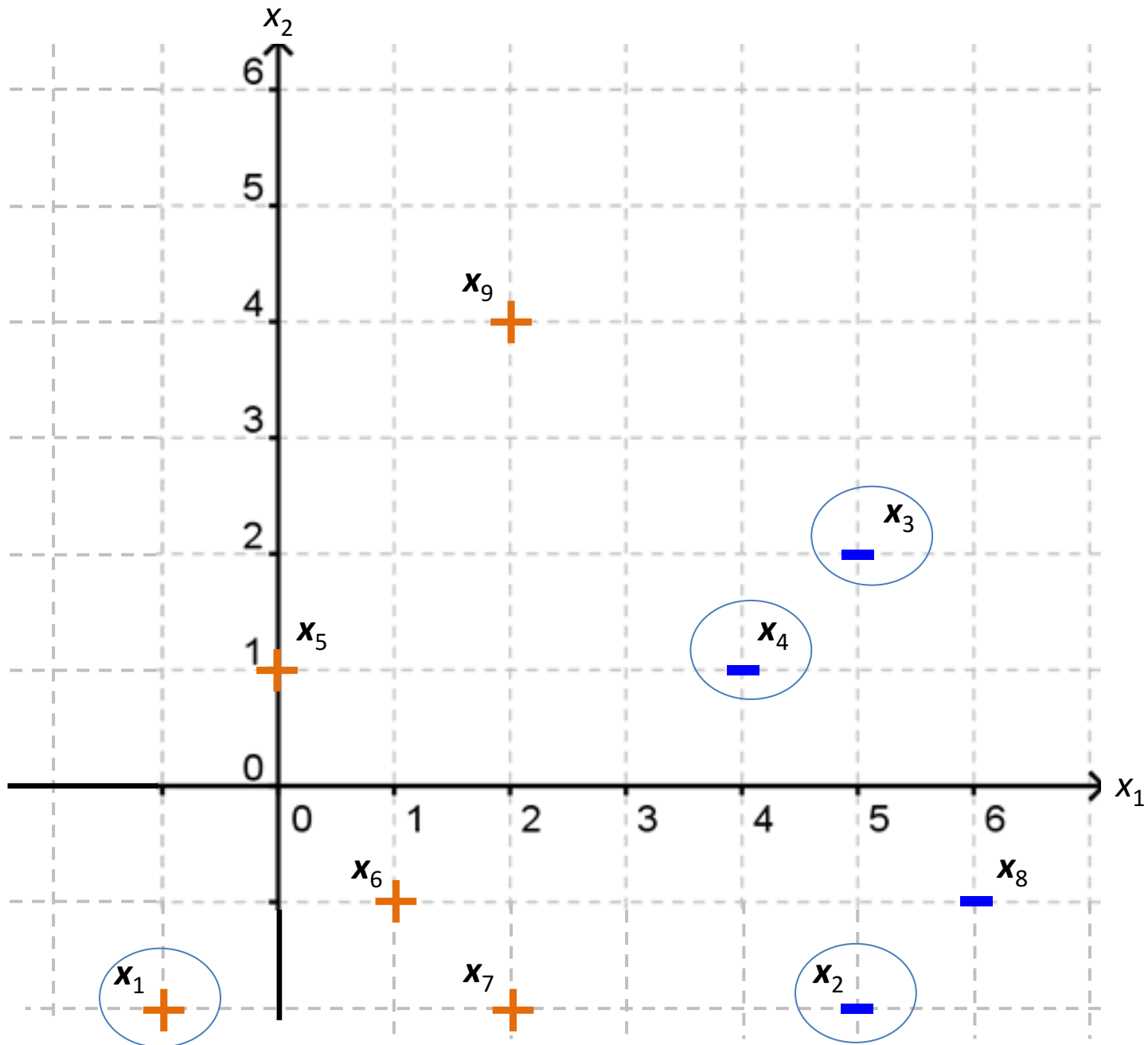
- if $\alpha_i = 0$ for $\vec{x}_i \in S$, discard \vec{x}_i

- add examples to S until $|S| = K$

Meta-optimization: example



Meta-optimization: example



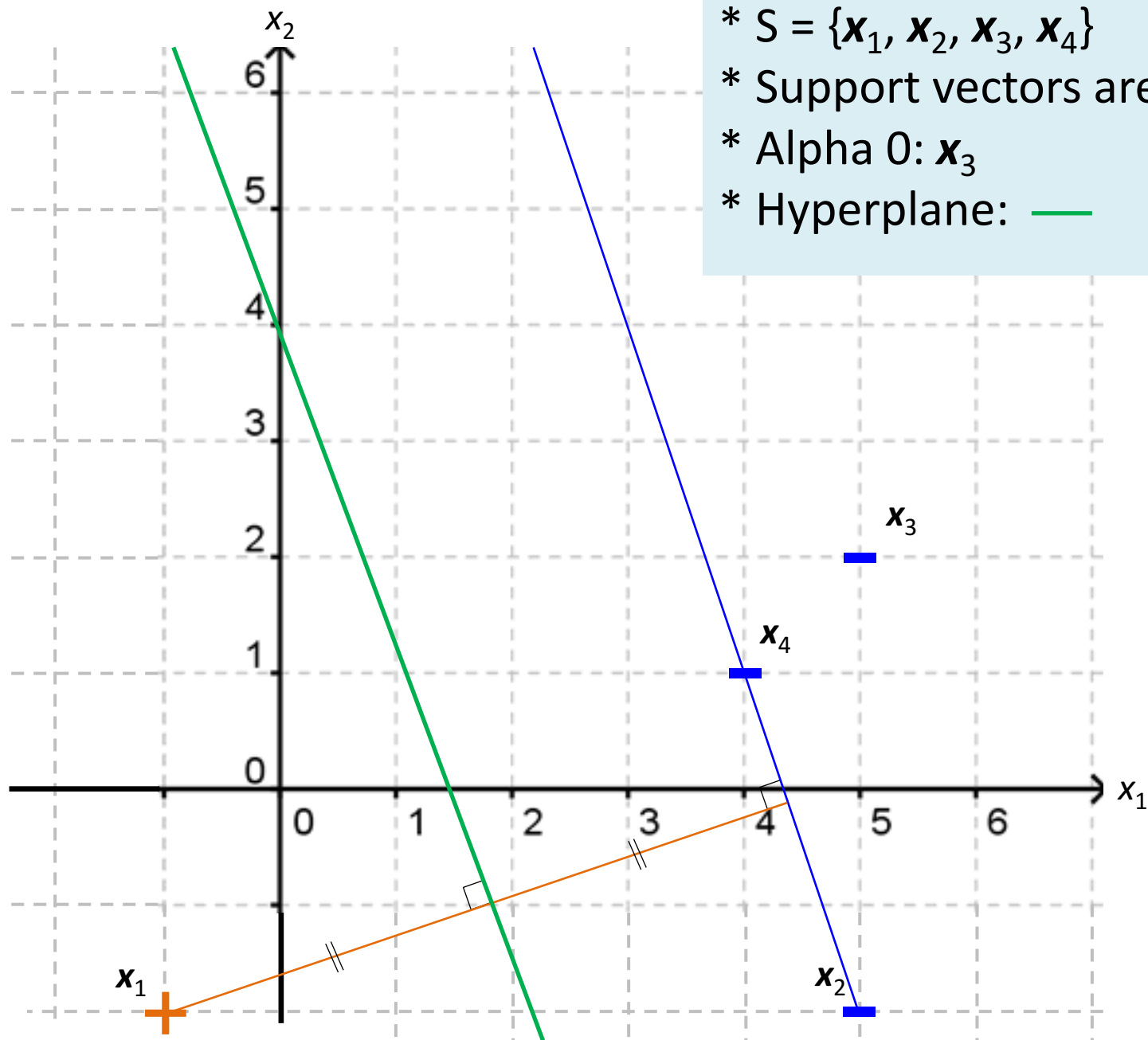
Round 1:

* $S = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$

* Support vectors are: $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4$

* Alpha 0: \mathbf{x}_3

* Hyperplane: —



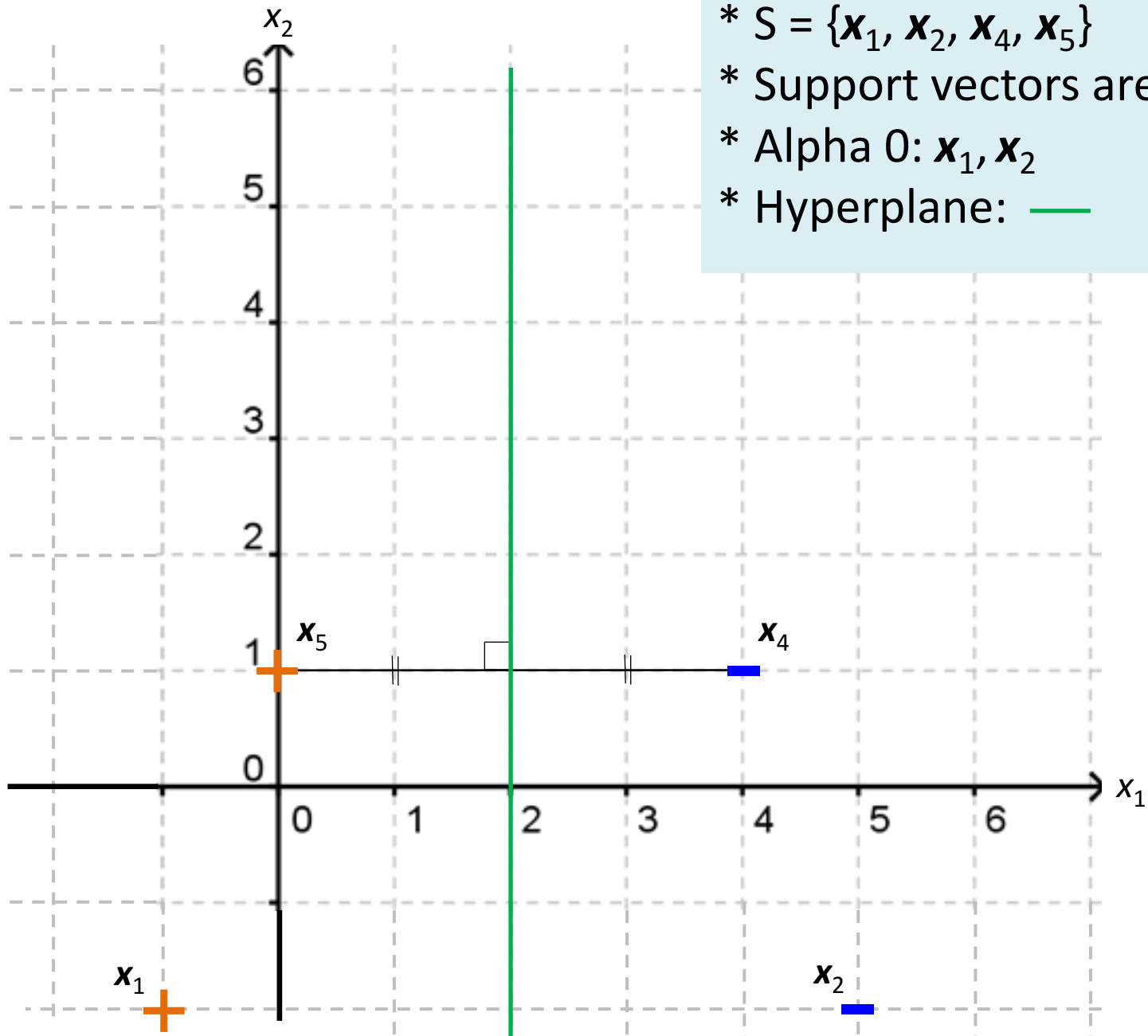
Round 1:

* $S = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5\}$

* Support vectors are: $\mathbf{x}_4, \mathbf{x}_5$

* Alpha 0: $\mathbf{x}_1, \mathbf{x}_2$

* Hyperplane: —



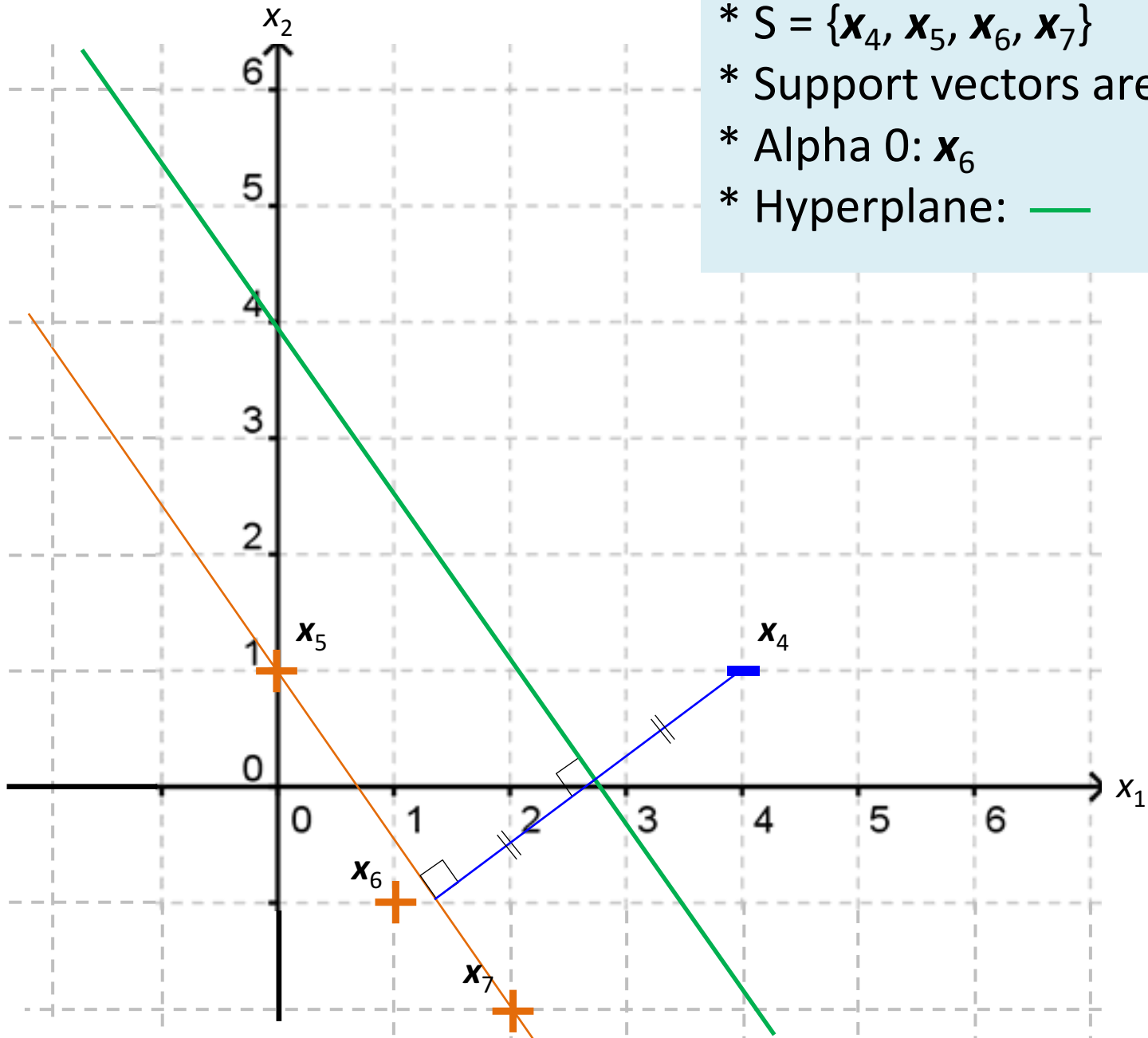
Round 3:

* $S = \{\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$

* Support vectors are: $\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_7$

* Alpha 0: \mathbf{x}_6

* Hyperplane: —



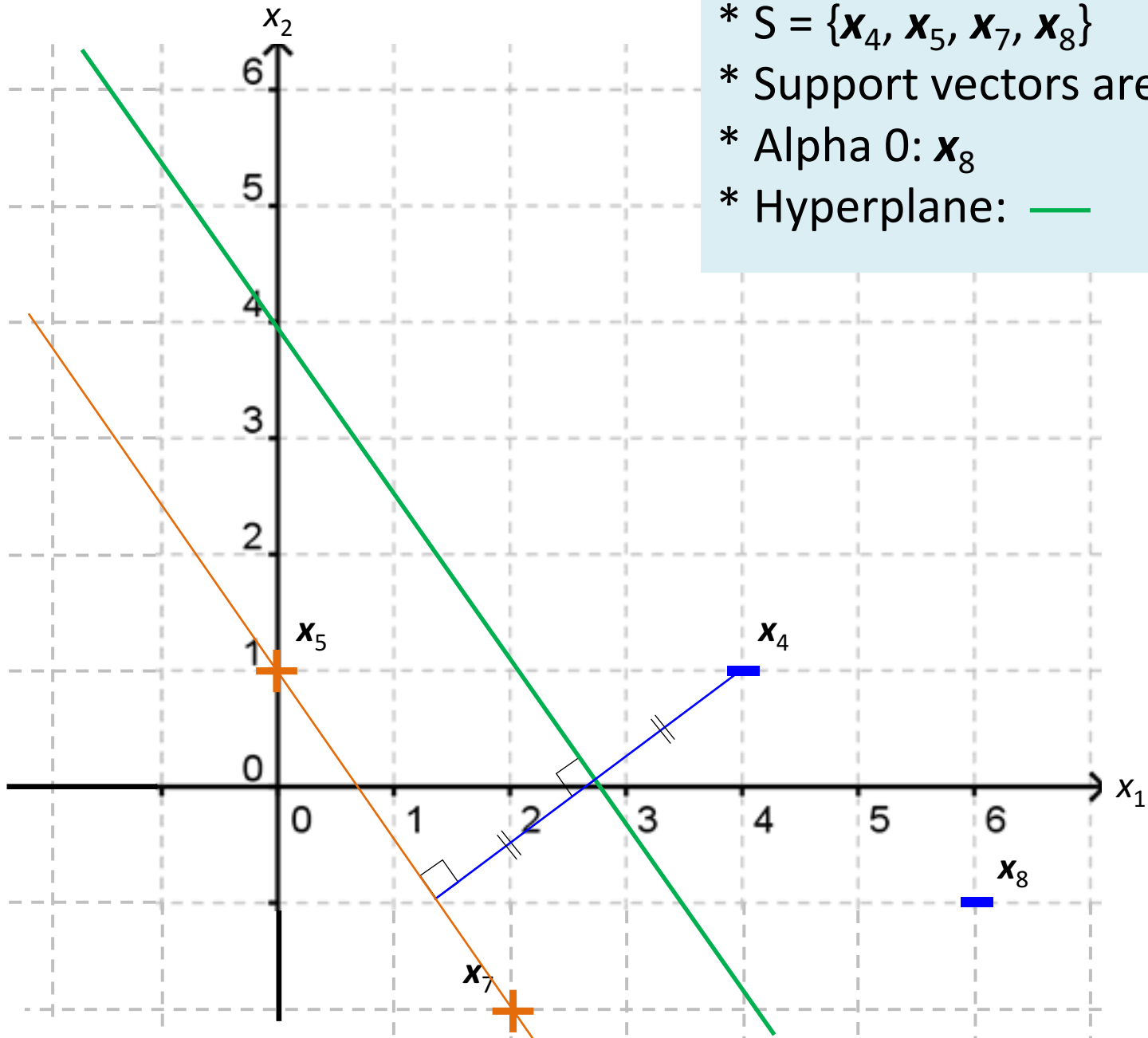
Round 4:

* $S = \{\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_7, \mathbf{x}_8\}$

* Support vectors are: $\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_7$

* Alpha 0: \mathbf{x}_8

* Hyperplane: —



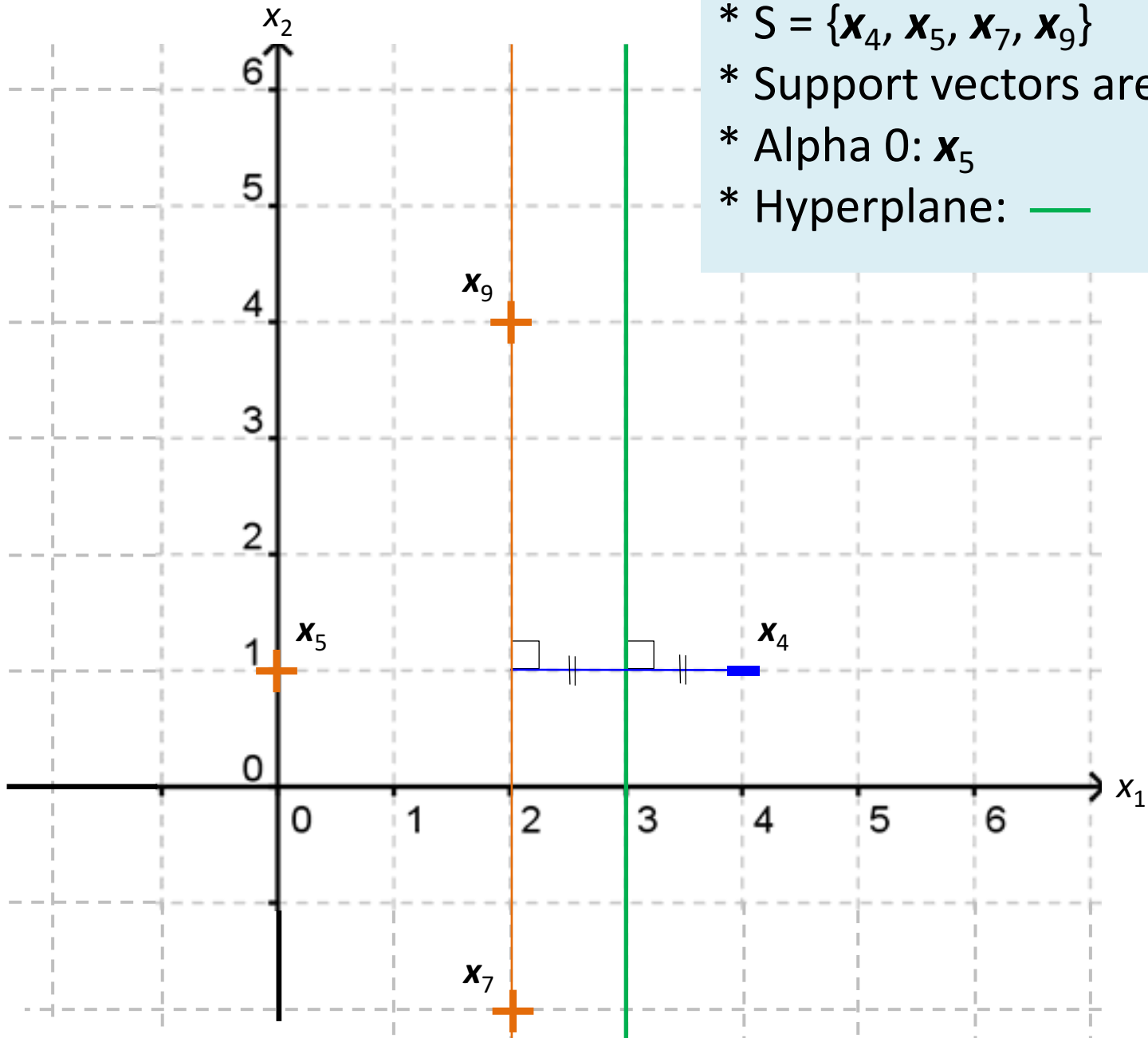
Round 5:

* $S = \{\mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_7, \mathbf{x}_9\}$

* Support vectors are: $\mathbf{x}_4, \mathbf{x}_7, \mathbf{x}_9$

* Alpha 0: \mathbf{x}_5

* Hyperplane: —



Handout 12, Final Solution

