

Perceptron

Machine Learning
CS4824/ECE4424
Bert Huang
Virginia Tech

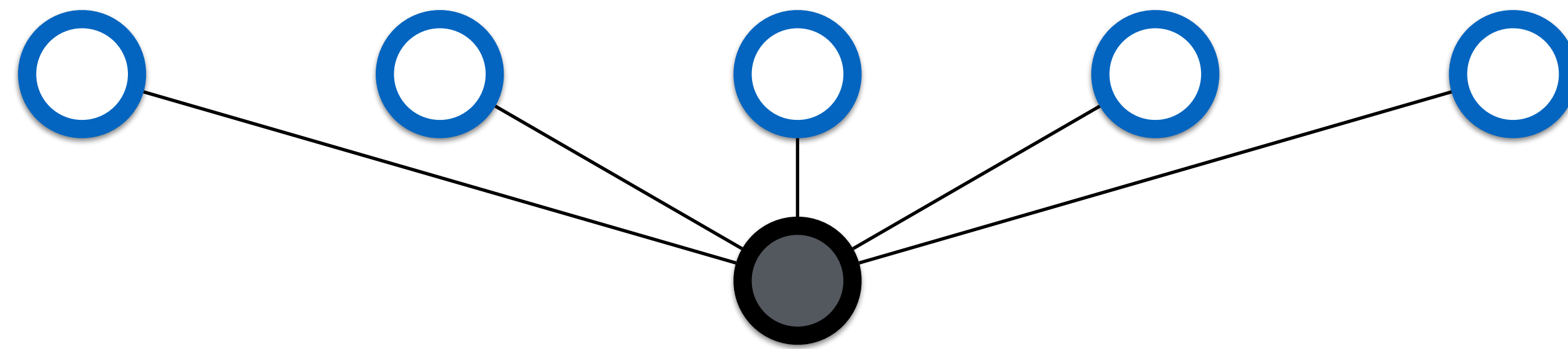
Plan

- Perceptron
- Multi-class perceptron
- HW1 Q/A

Perceptron

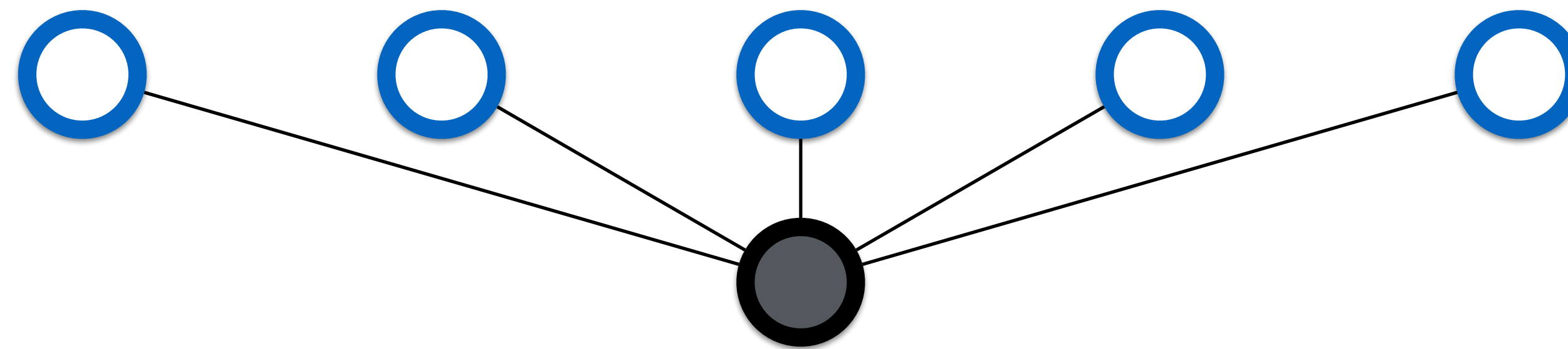
Perceptron

- AKA single-layer neural network



Perceptron

- AKA single-layer neural network

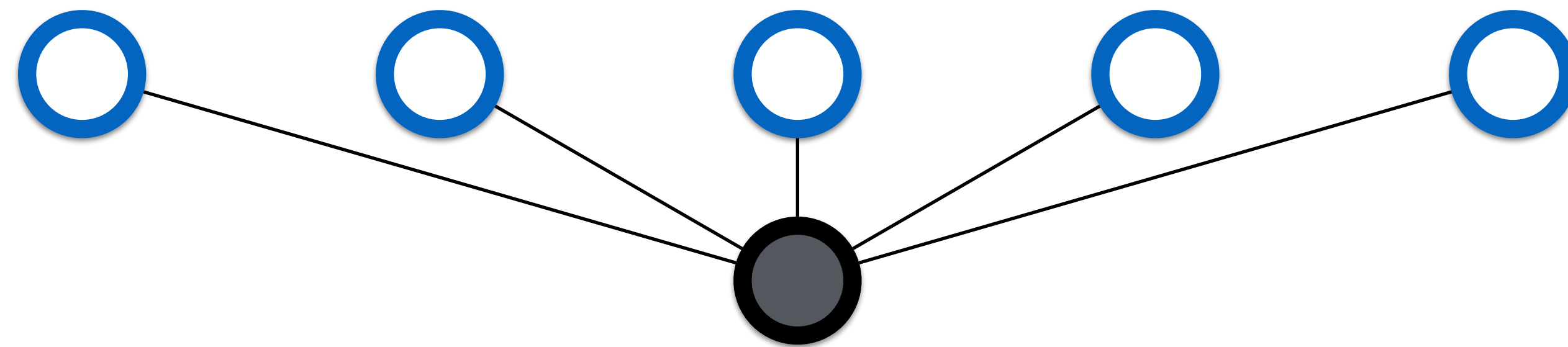


- linear classifier

$$f(x) = \text{sign} \left(\sum_i w_i x_i \right)$$

Perceptron

- AKA single-layer neural network

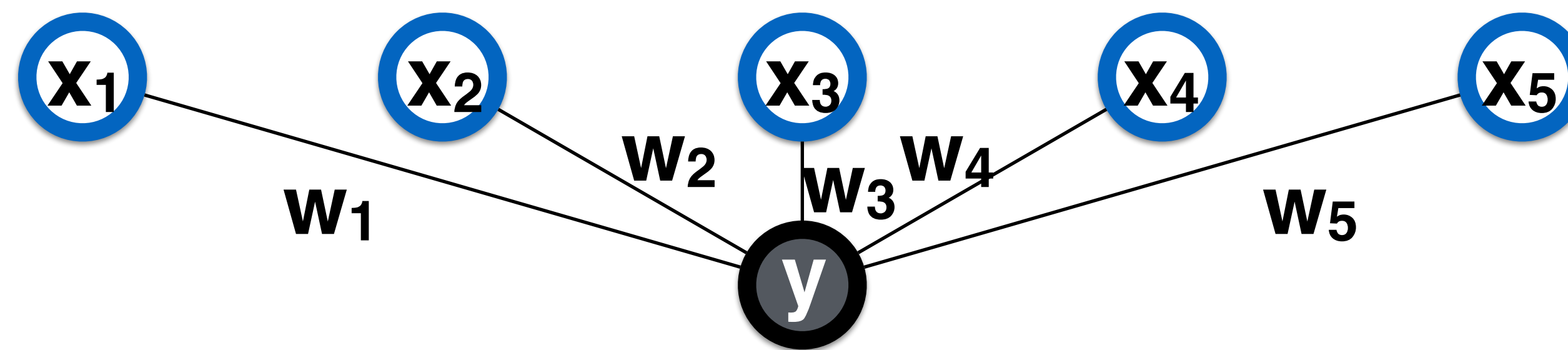


- linear classifier

$$f(x) = \text{sign} \left(\sum_i w_i x_i \right) = \begin{cases} +1 & \text{if } \sum_i w_i x_i \geq 0 \\ -1 & \text{if } \sum_i w_i x_i < 0 \end{cases}$$

Perceptron

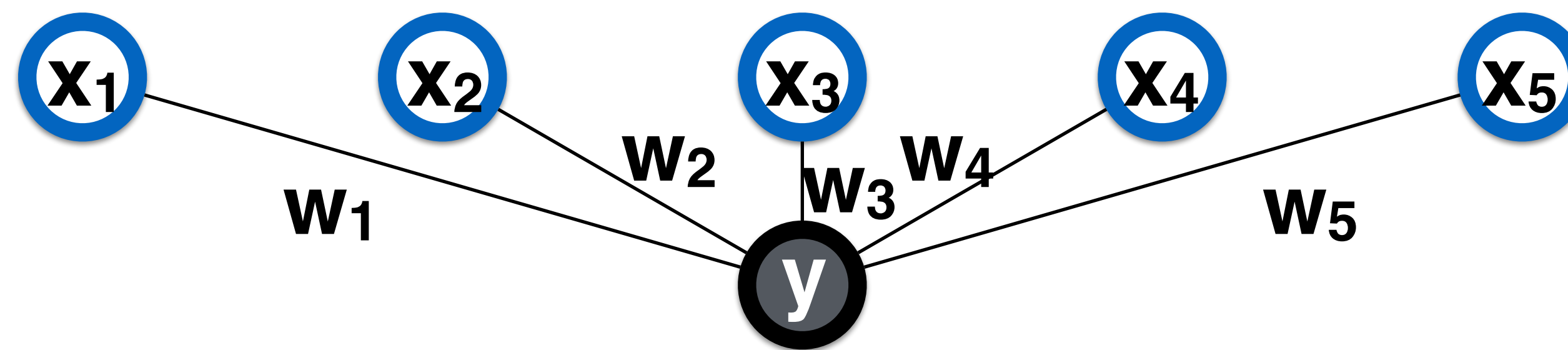
- AKA single-layer neural network



- linear classifier
- $$f(x) = \text{sign} \left(\sum_i w_i x_i \right) = \begin{cases} +1 & \text{if } \sum_i w_i x_i \geq 0 \\ -1 & \text{if } \sum_i w_i x_i < 0 \end{cases}$$

Perceptron

- AKA single-layer neural network



- linear classifier
- $$f(x) = \text{sign} \left(\sum_i w_i x_i \right) = \begin{cases} +1 & \text{if } \sum_i w_i x_i \geq 0 \\ -1 & \text{if } \sum_i w_i x_i < 0 \end{cases}$$
- online learner

Online Training

- Get example $\mathbf{x}$ and oracle label $\mathbf{y}$
- Try current classifier $\mathbf{f}(\mathbf{x})$
- If $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, celebrate
- If $\mathbf{f}(\mathbf{x}) \neq \mathbf{y}$, fix it

Online Training

- Get example $\mathbf{x}$ and oracle label $\mathbf{y}$
- Try current classifier $\mathbf{f}(\mathbf{x})$ $\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w}_1\mathbf{x}_1 + \mathbf{w}_2\mathbf{x}_2 + \mathbf{w}_3\mathbf{x}_4 + \dots)$
- If $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, celebrate
- If $\mathbf{f}(\mathbf{x}) \neq \mathbf{y}$, fix it

Online Training

- Get example $\mathbf{x}$ and oracle label $\mathbf{y}$
- Try current classifier $\mathbf{f}(\mathbf{x})$ $\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w}_1\mathbf{x}_1 + \mathbf{w}_2\mathbf{x}_2 + \mathbf{w}_3\mathbf{x}_4 + \dots)$
- If $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, celebrate if $\mathbf{y} = +1$ and $\mathbf{f}(\mathbf{x}) = -1$, $\mathbf{w} \cdot \mathbf{x}$ is too small
- If $\mathbf{f}(\mathbf{x}) \neq \mathbf{y}$, fix it

Online Training

- Get example $\mathbf{x}$ and oracle label $\mathbf{y}$
- Try current classifier $\mathbf{f}(\mathbf{x})$ $\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w}_1\mathbf{x}_1 + \mathbf{w}_2\mathbf{x}_2 + \mathbf{w}_3\mathbf{x}_4 + \dots)$
- If $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, celebrate if $\mathbf{y} = +1$ and $\mathbf{f}(\mathbf{x}) = -1$, $\mathbf{w} \cdot \mathbf{x}$ is too small
- If $\mathbf{f}(\mathbf{x}) \neq \mathbf{y}$, fix it if $\mathbf{y} = -1$ and $\mathbf{f}(\mathbf{x}) = +1$, $\mathbf{w} \cdot \mathbf{x}$ is too big

How to Fix Your Broken Perceptron

How to Fix Your Broken Perceptron

$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$

How to Fix Your Broken Perceptron

$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$

$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$

How to Fix Your Broken Perceptron

$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$

$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$

$w \leftarrow w$ if $f(x) = y$

How to Fix Your Broken Perceptron

$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$

$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$f(x) = \text{sign}(w_1 x_1 + w_2 x_2) = +1$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$f(x) = \text{sign}(w_1 x_1 + w_2 x_2) = +1$
 $y = -1$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$$f(x) = \text{sign}(w_1 x_1 + w_2 x_2) = +1$$
$$y = -1$$

$$2(3) + 1(1)$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$$f(x) = \text{sign}(w_1 x_1 + w_2 x_2) = +1$$
$$y = -1$$

$$2(3) + 1(1)$$

$$W_1 = 2 - 3$$

$$W_2 = 1 - 1$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$$f(x) = \text{sign}(w_1 x_1 + w_2 x_2) = +1$$
$$y = -1$$

$$2(3) + 1(1)$$

$$W_1 = 2 - 3$$

$$W_2 = 1 - 1$$

$$-1(3) + 0(1)$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$$y = +1, f(x) = -1$$
$$w \cdot x$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$y = +1, f(x) = -1$

$w \cdot x$

$(w + x) \cdot x$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$y = +1, f(x) = -1$

$$w \cdot x \leq (w + x) \cdot x$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$y = +1, f(x) = -1$

$$w \cdot x \leq (w + x) \cdot x$$

$$w \cdot x + x \cdot x$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$y = +1, f(x) = -1$

$$w \cdot x \leq (w + x) \cdot x$$

$$w \cdot x + x \cdot x$$

$y = -1, f(x) = +1$

$$w \cdot x \geq (w - x) \cdot x$$

How to Fix Your Broken Perceptron

~~$w \leftarrow w - x$ if $f(x) = +1$ and $y = -1$~~

~~$w \leftarrow w + x$ if $f(x) = -1$ and $y = +1$~~

$w \leftarrow w$ if $f(x) = y$

$w \leftarrow w + yx$ if $f(x) \neq y$

$y = +1, f(x) = -1$

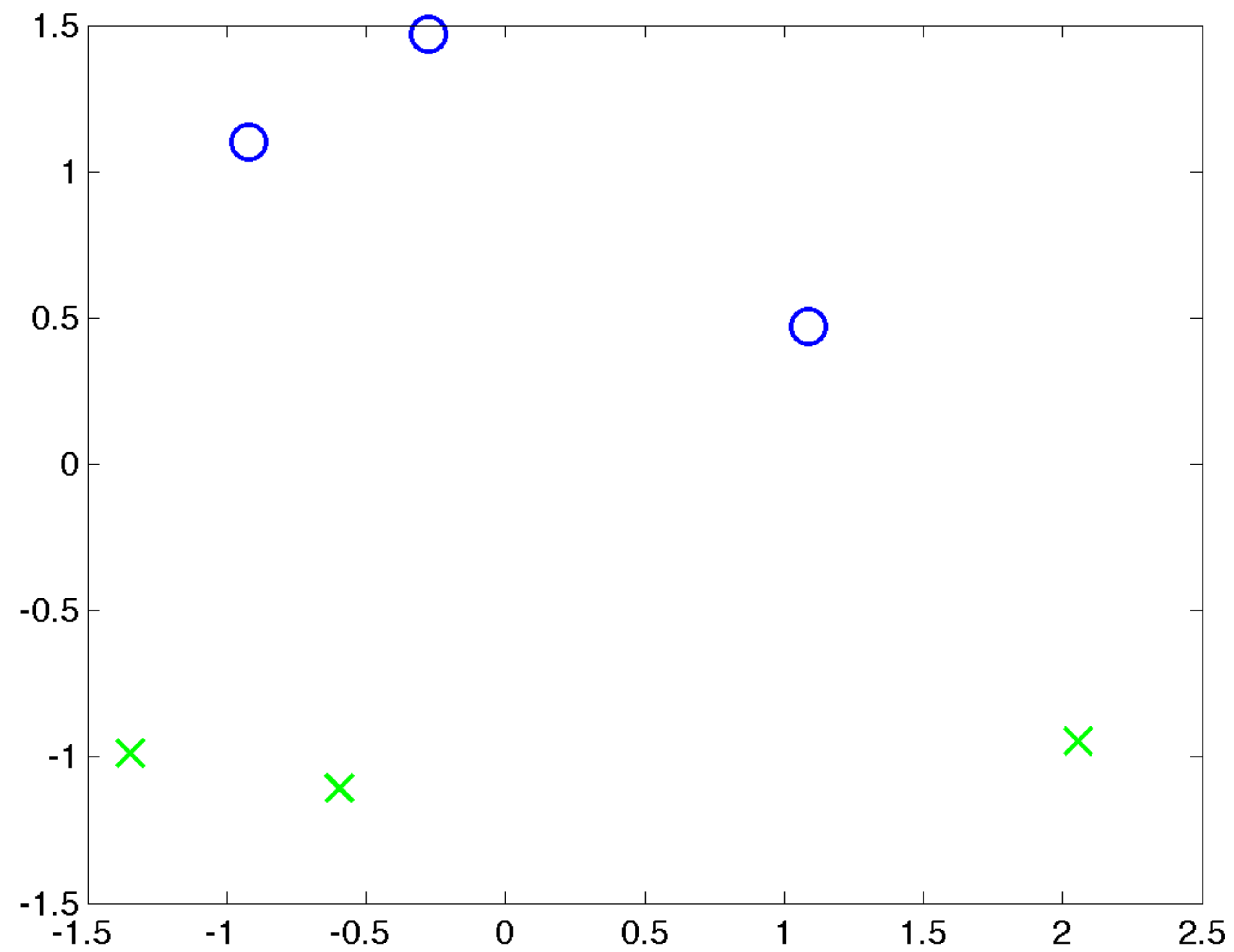
$$w \cdot x \leq (w + x) \cdot x$$

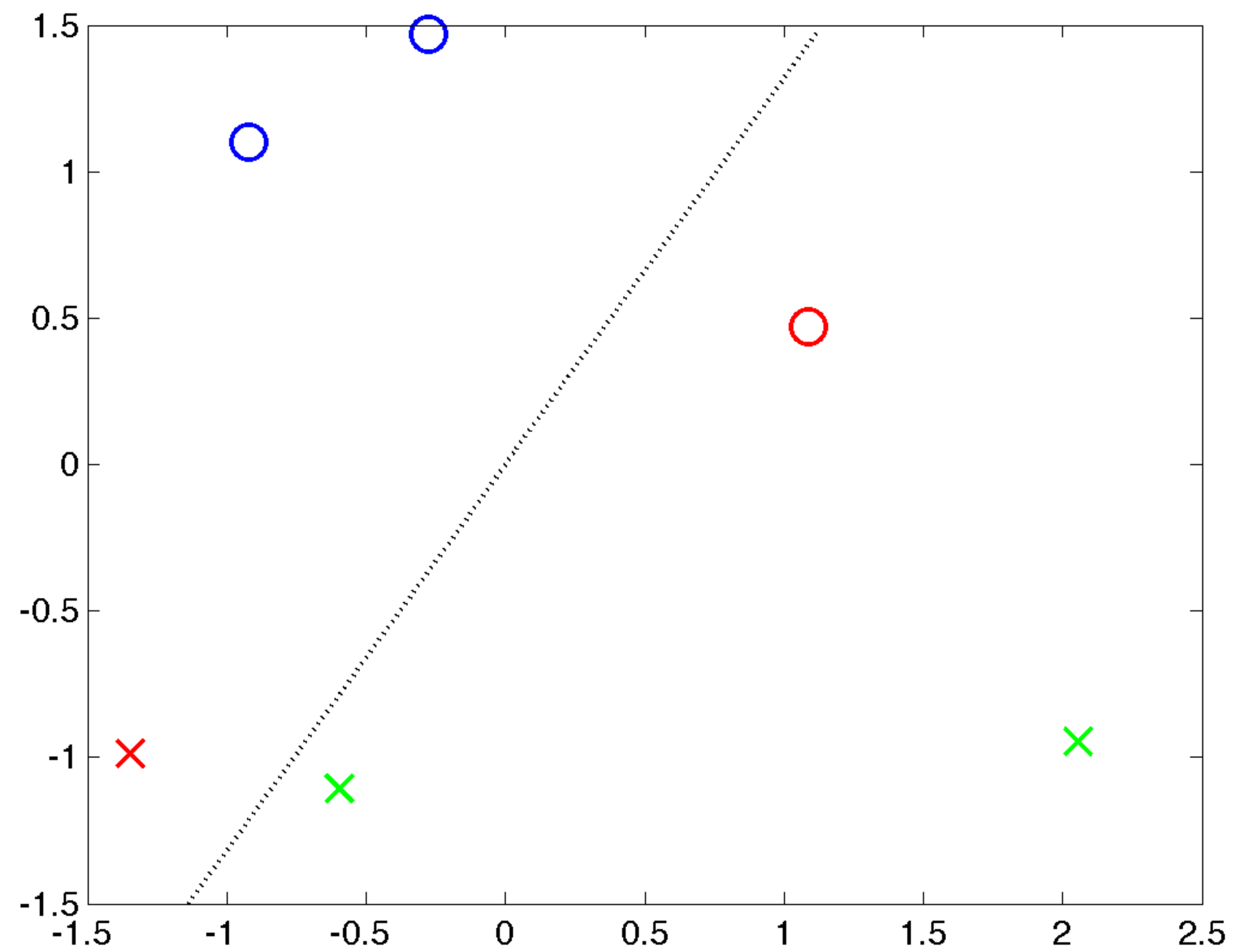
$$w \cdot x + x \cdot x$$

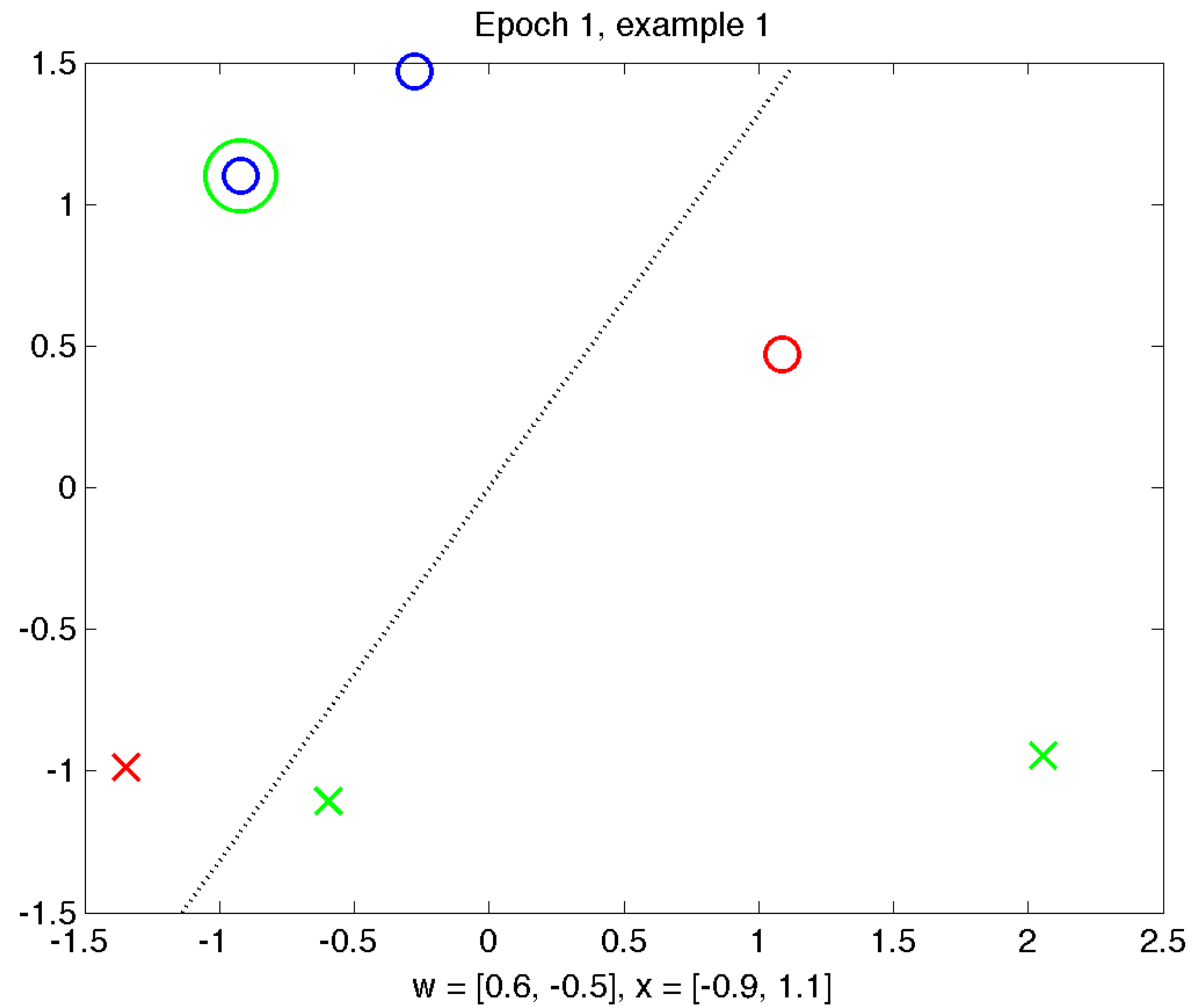
$y = -1, f(x) = +1$

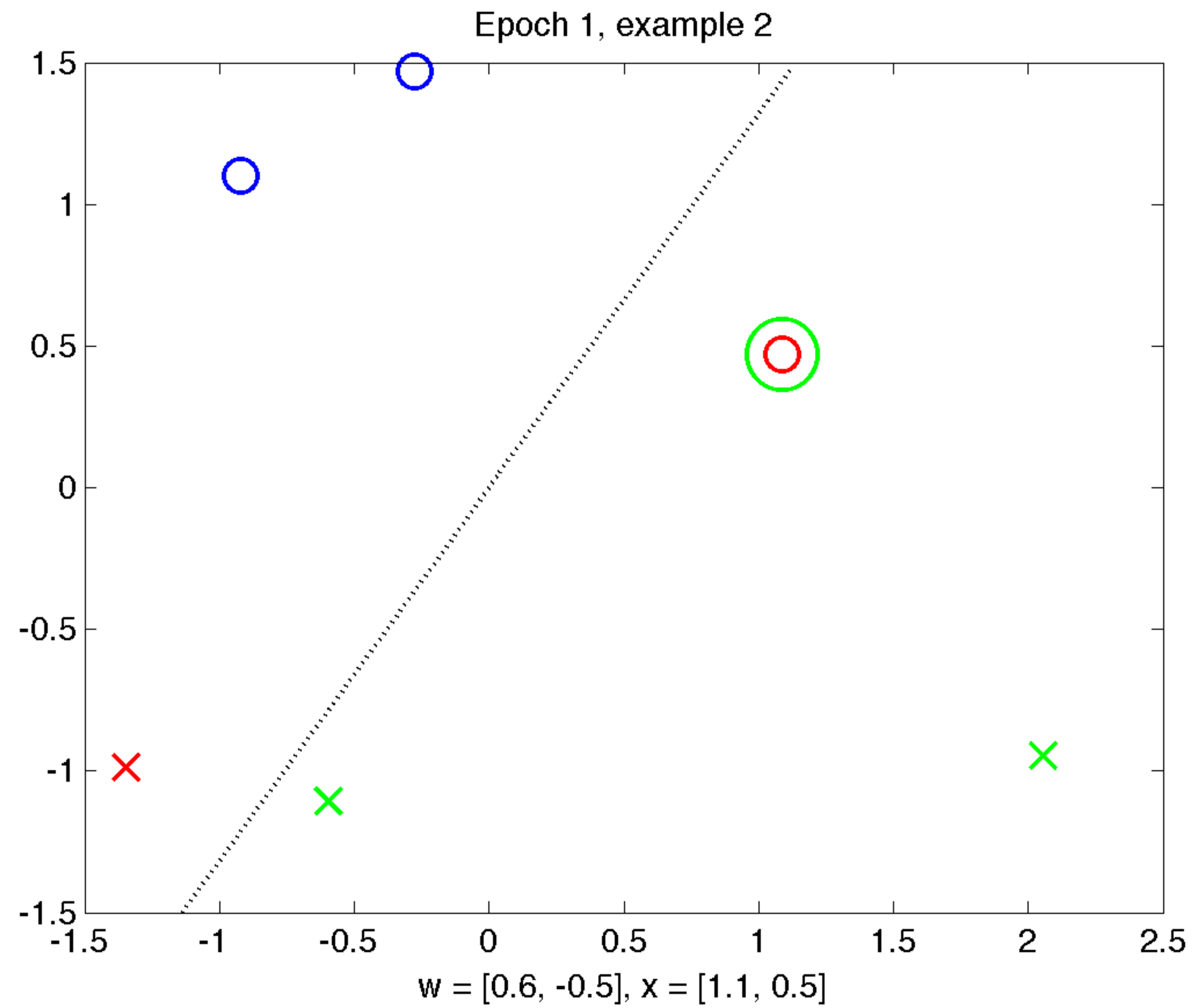
$$w \cdot x \geq (w - x) \cdot x$$

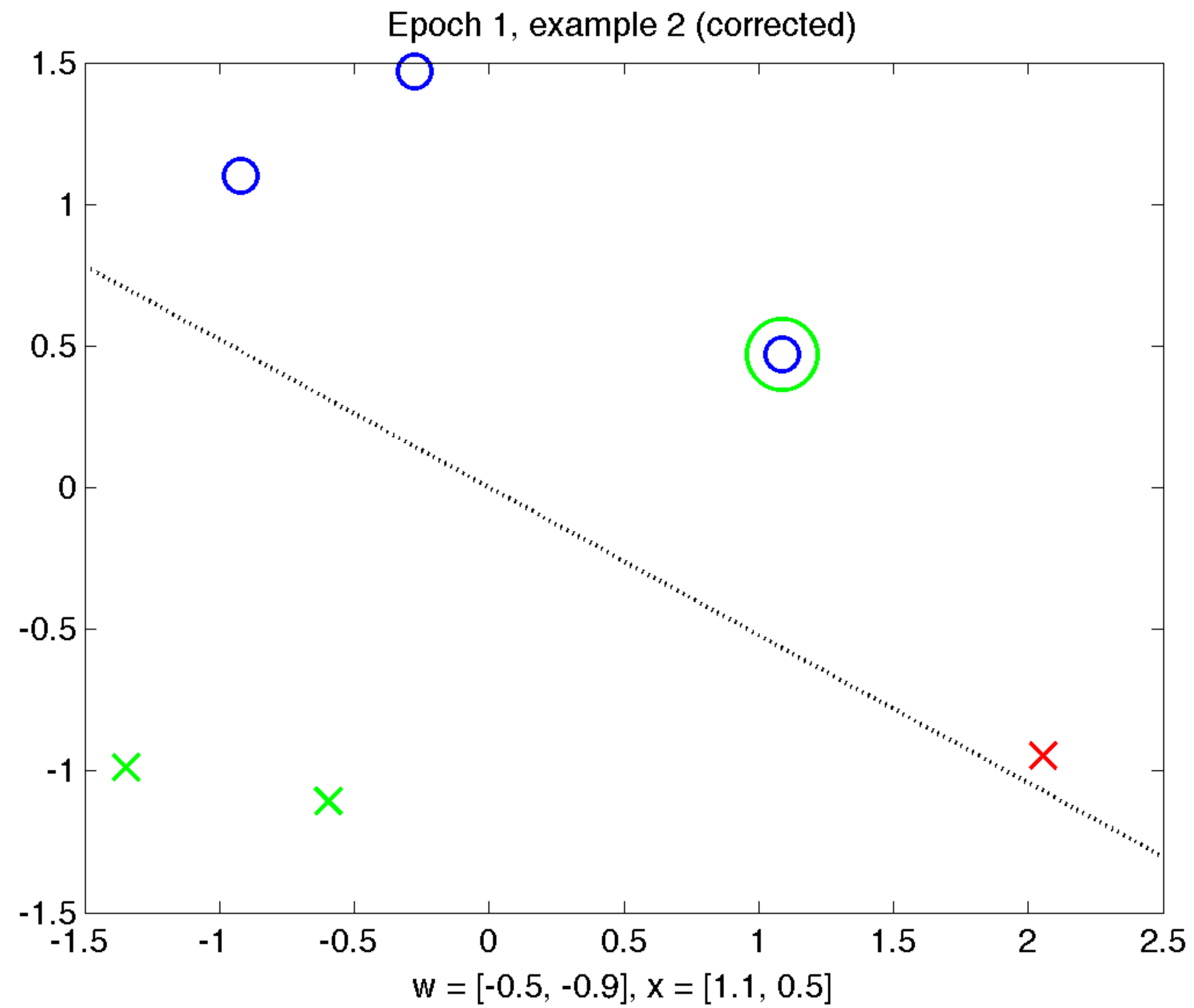
$$w \cdot x - x \cdot x$$

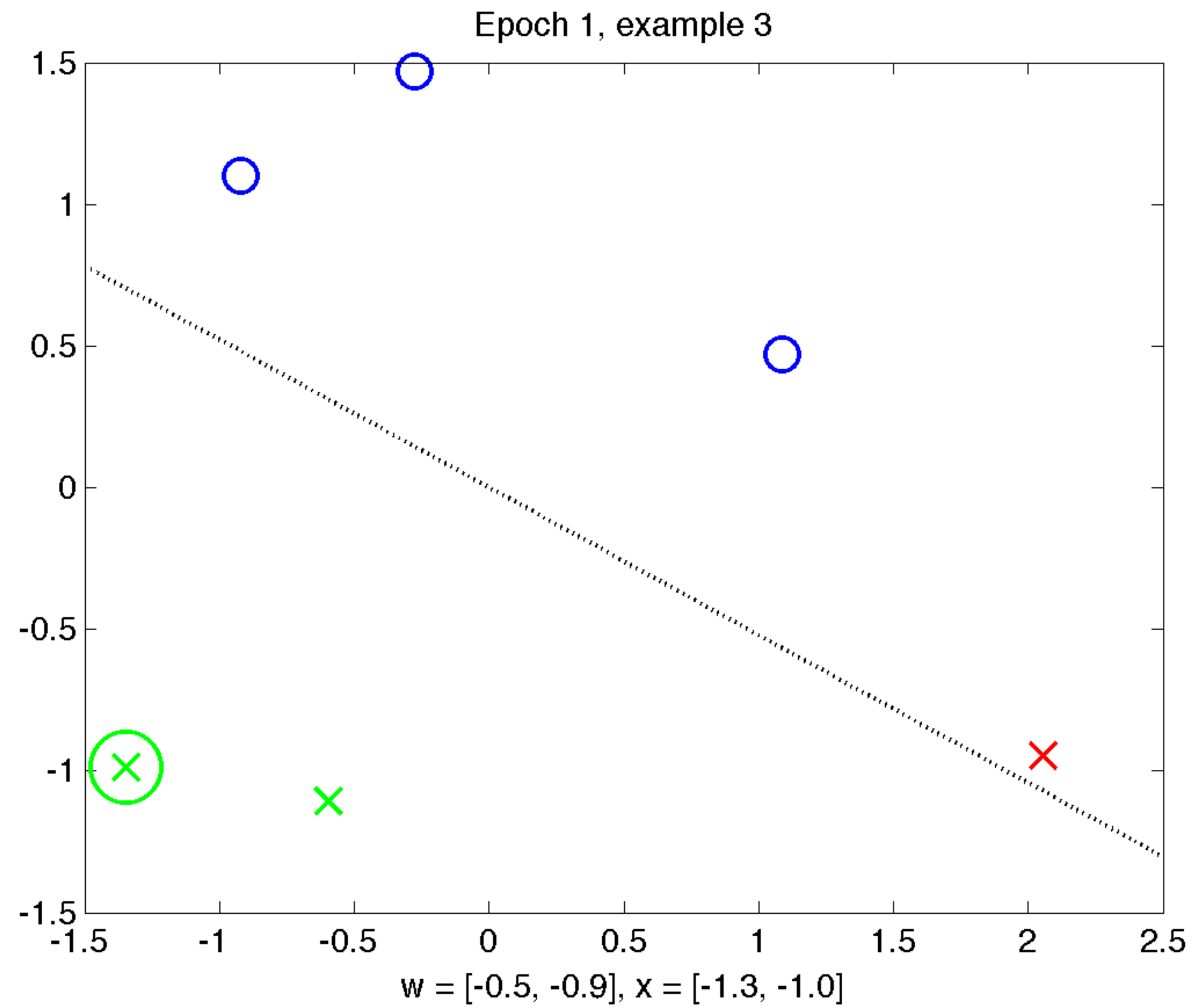


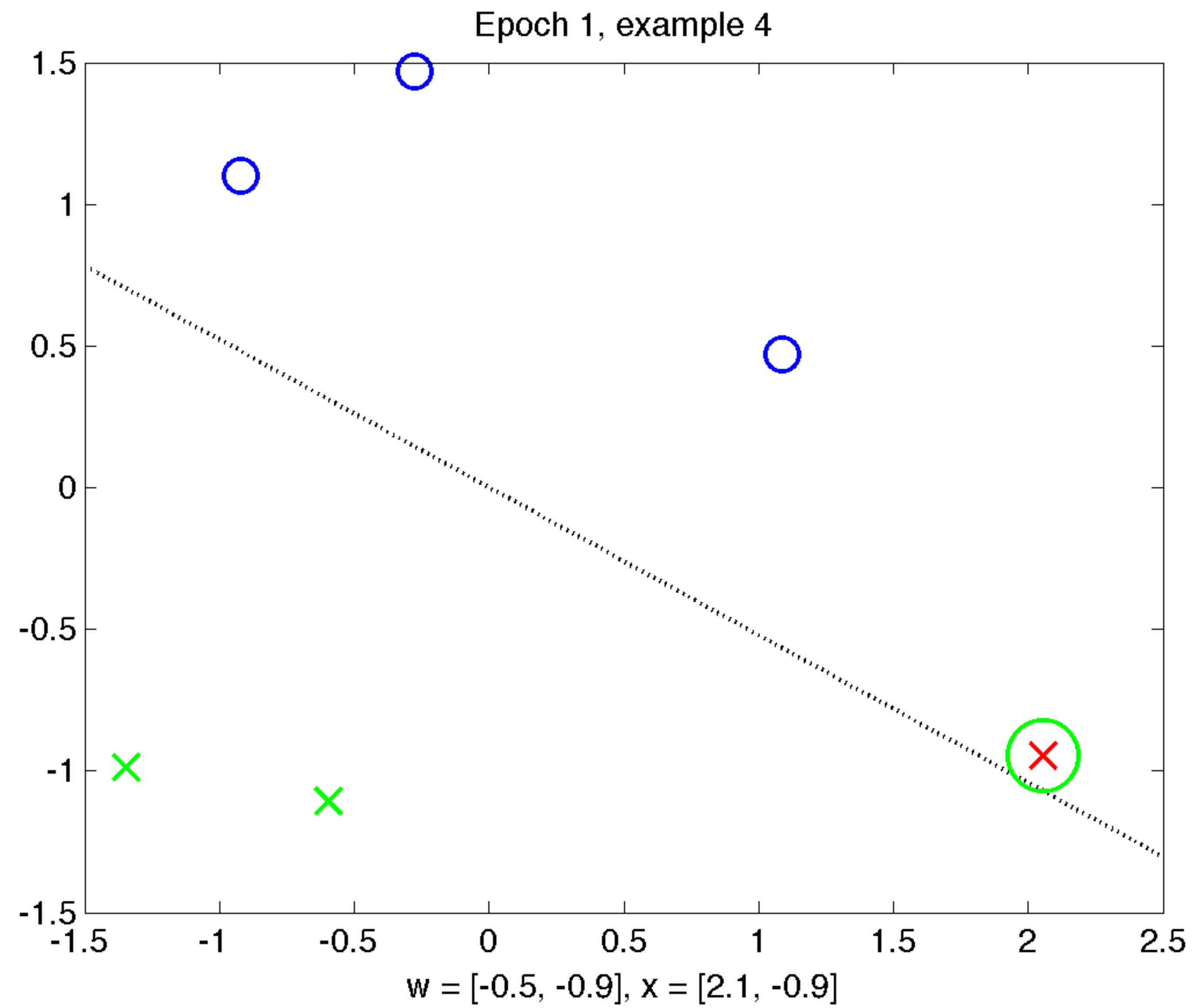


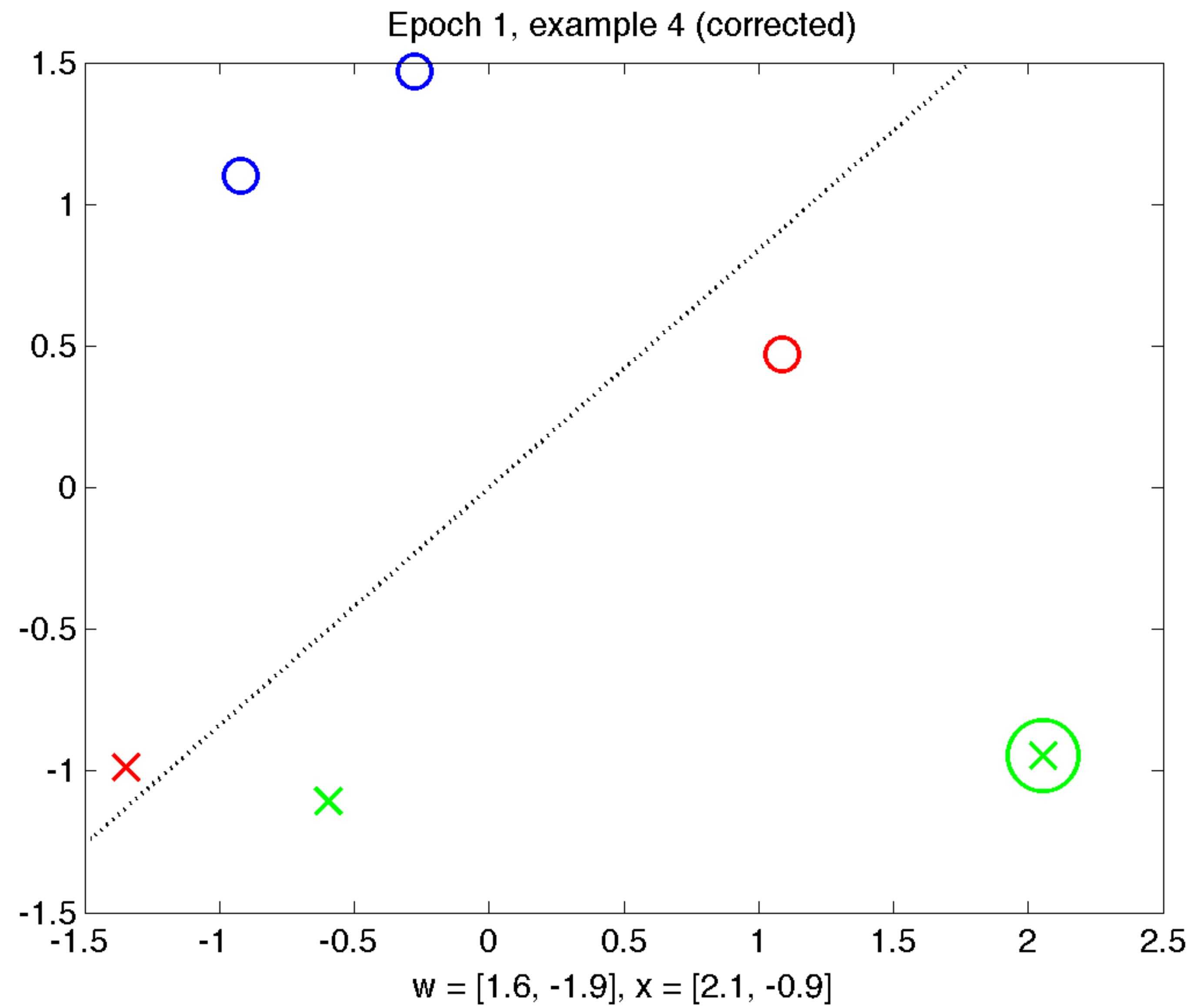


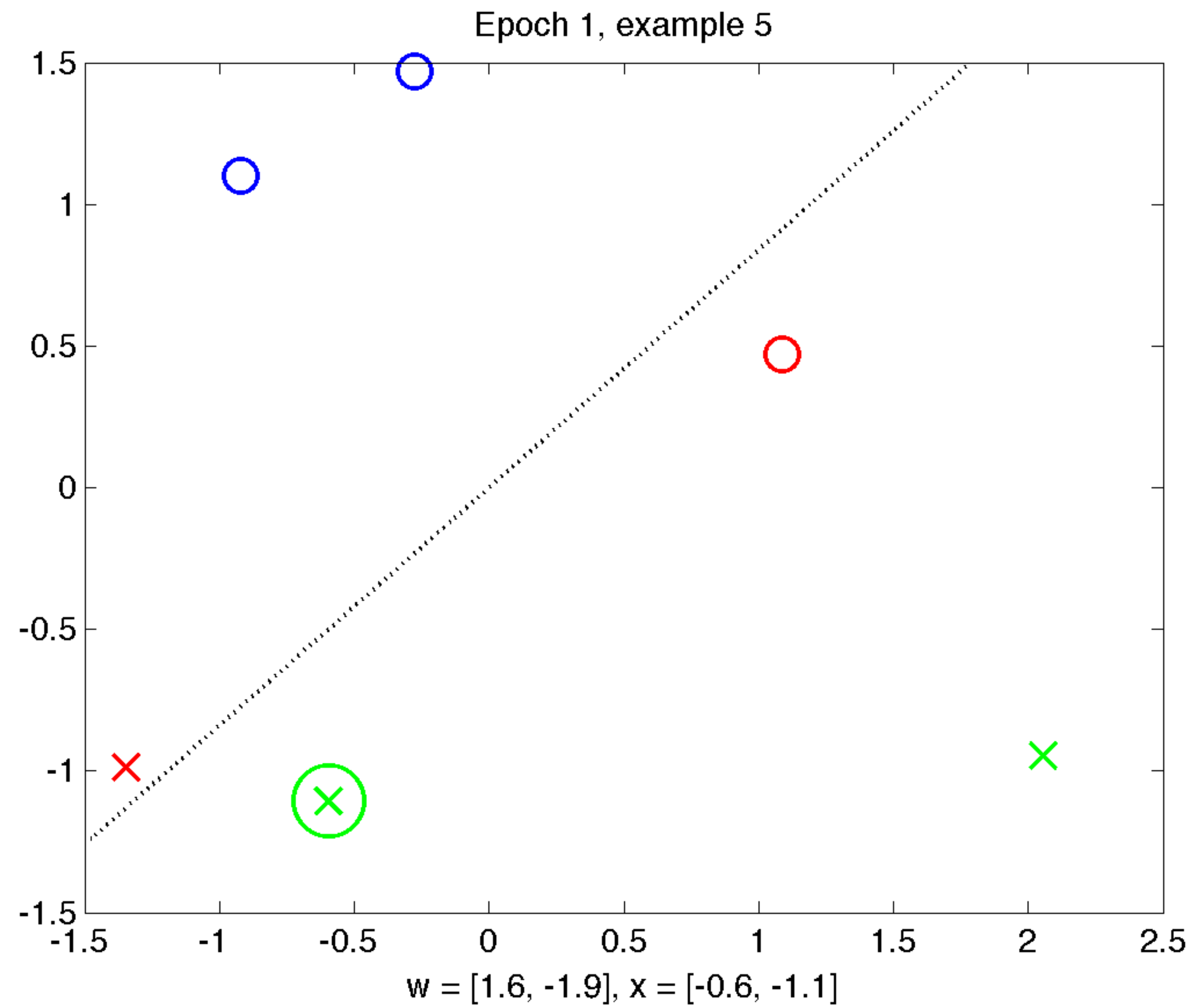


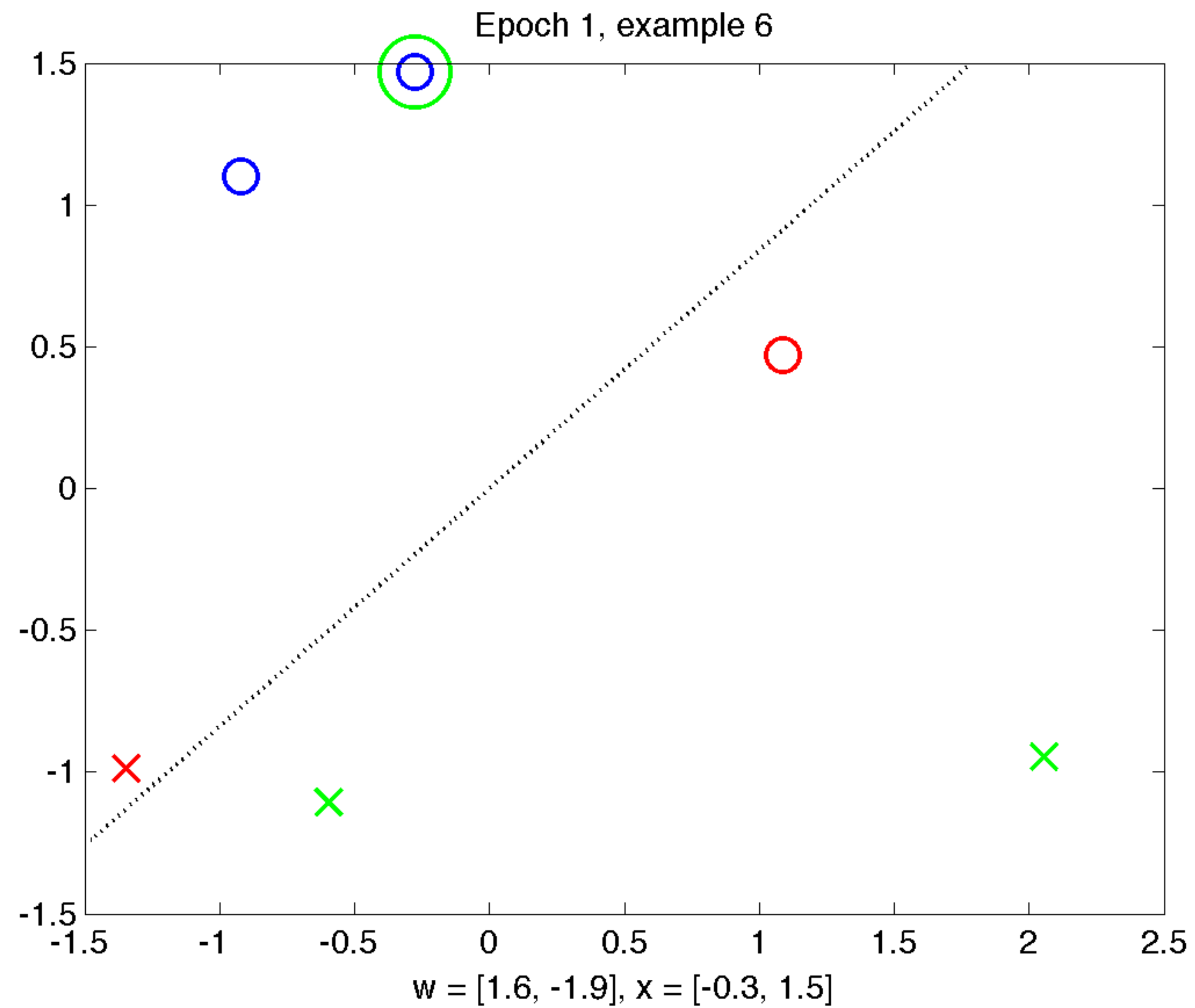


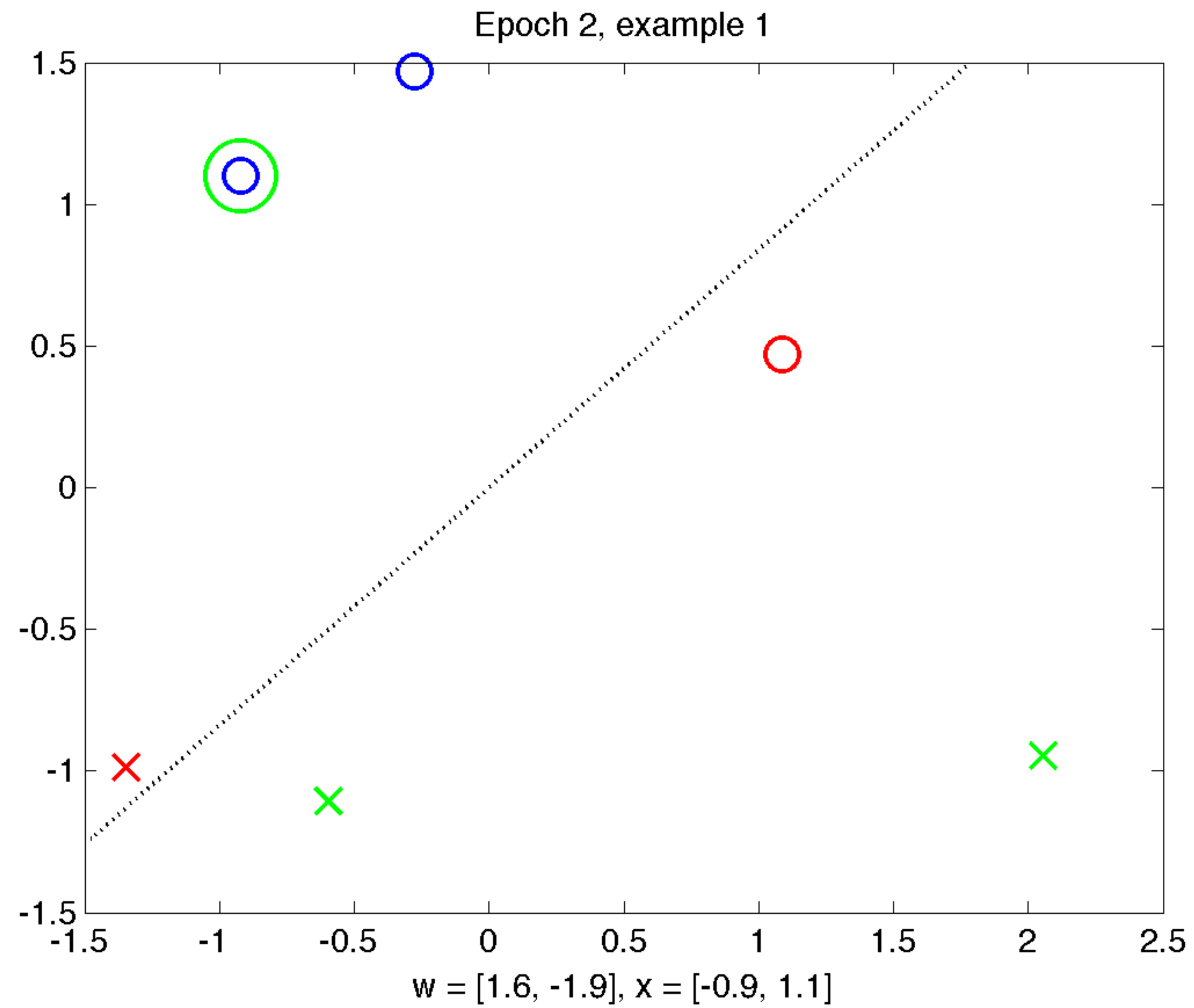


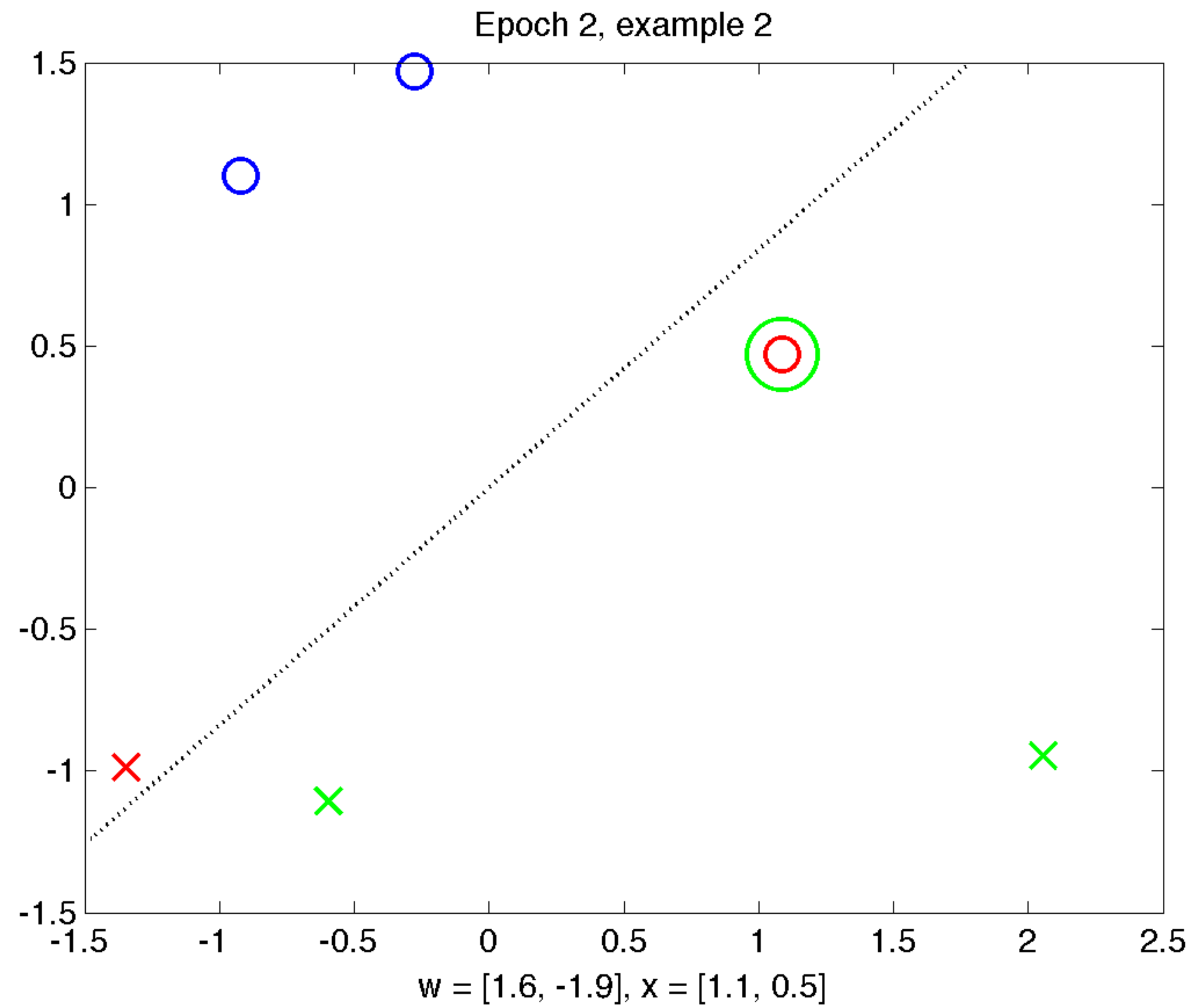


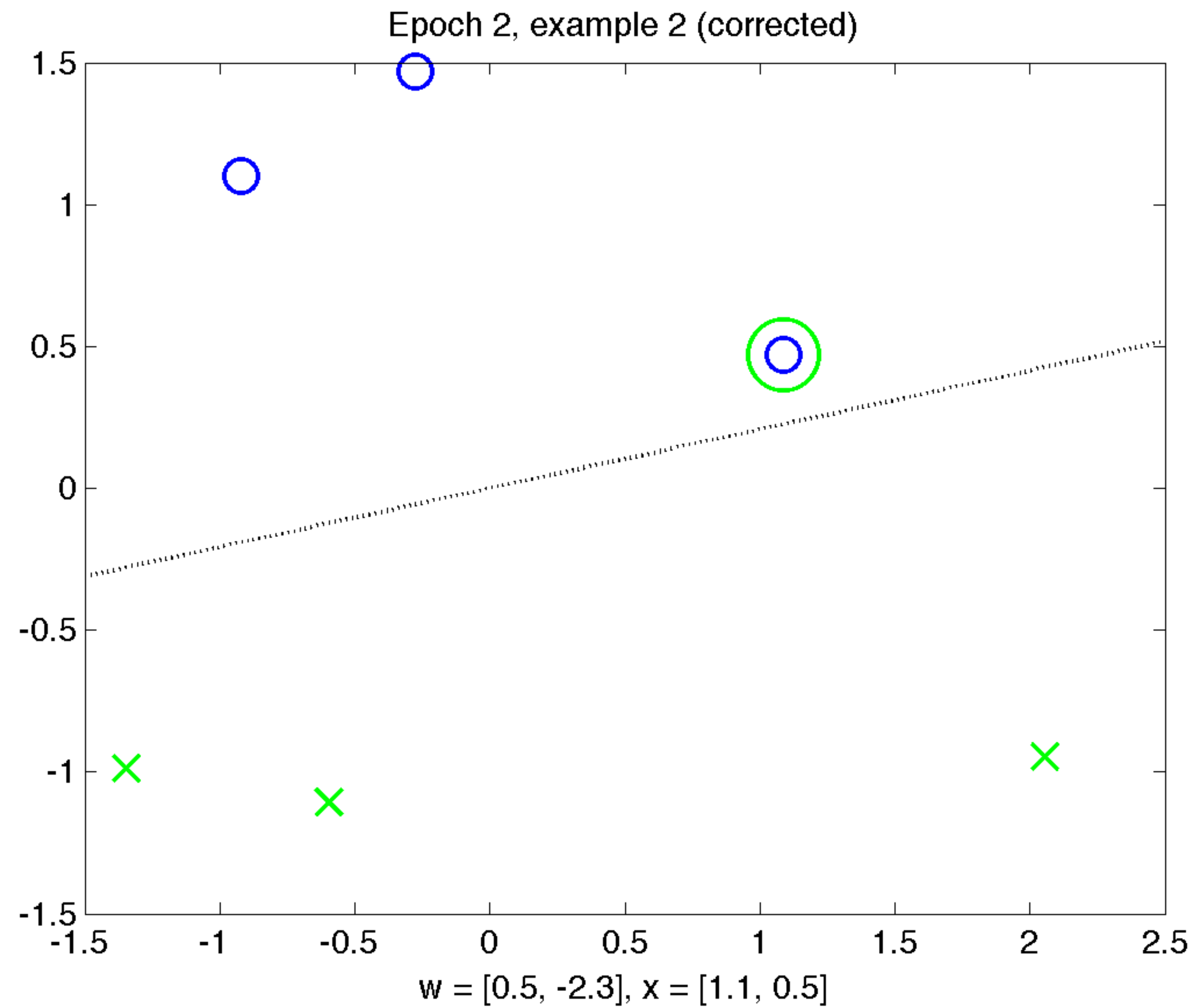


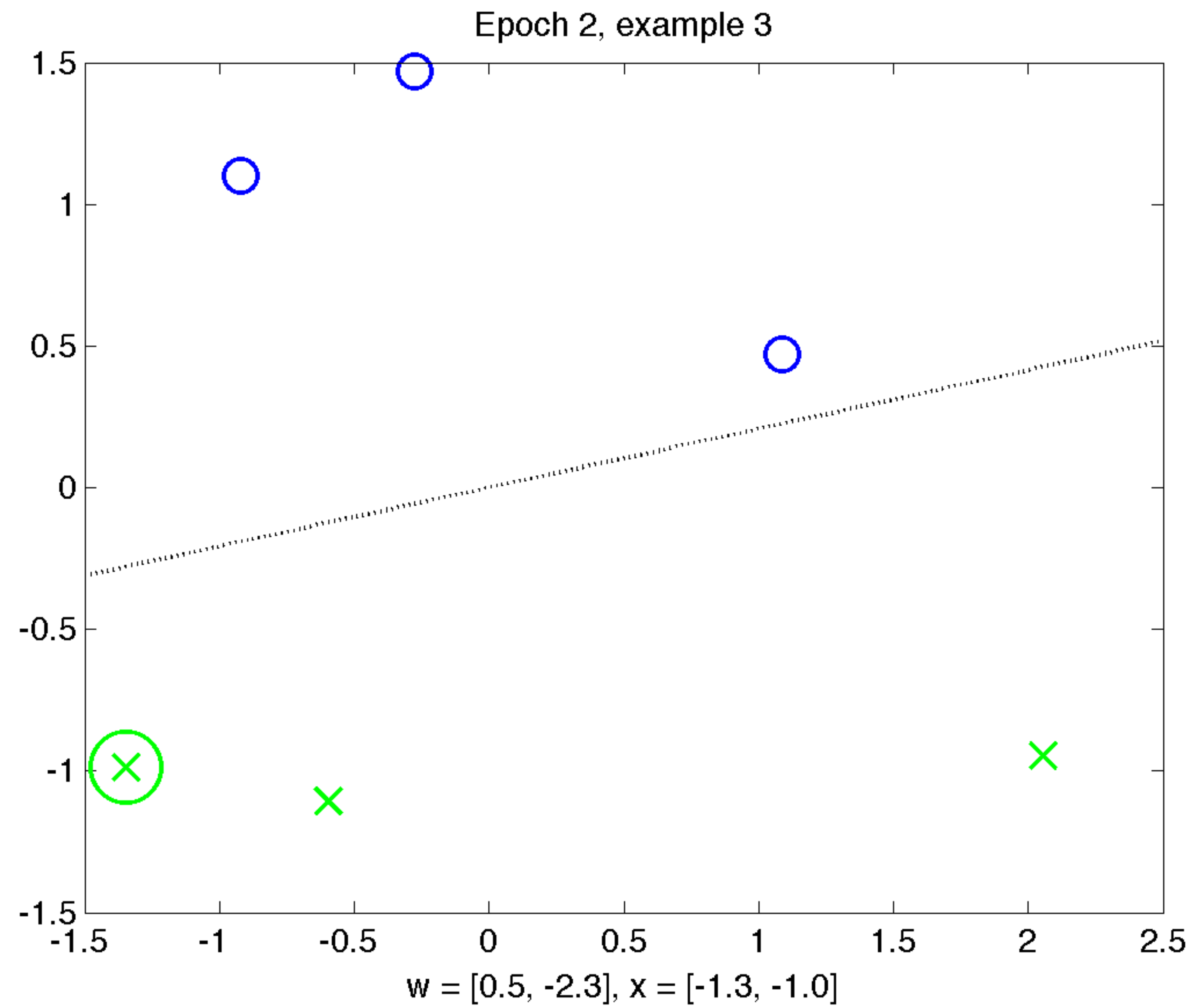


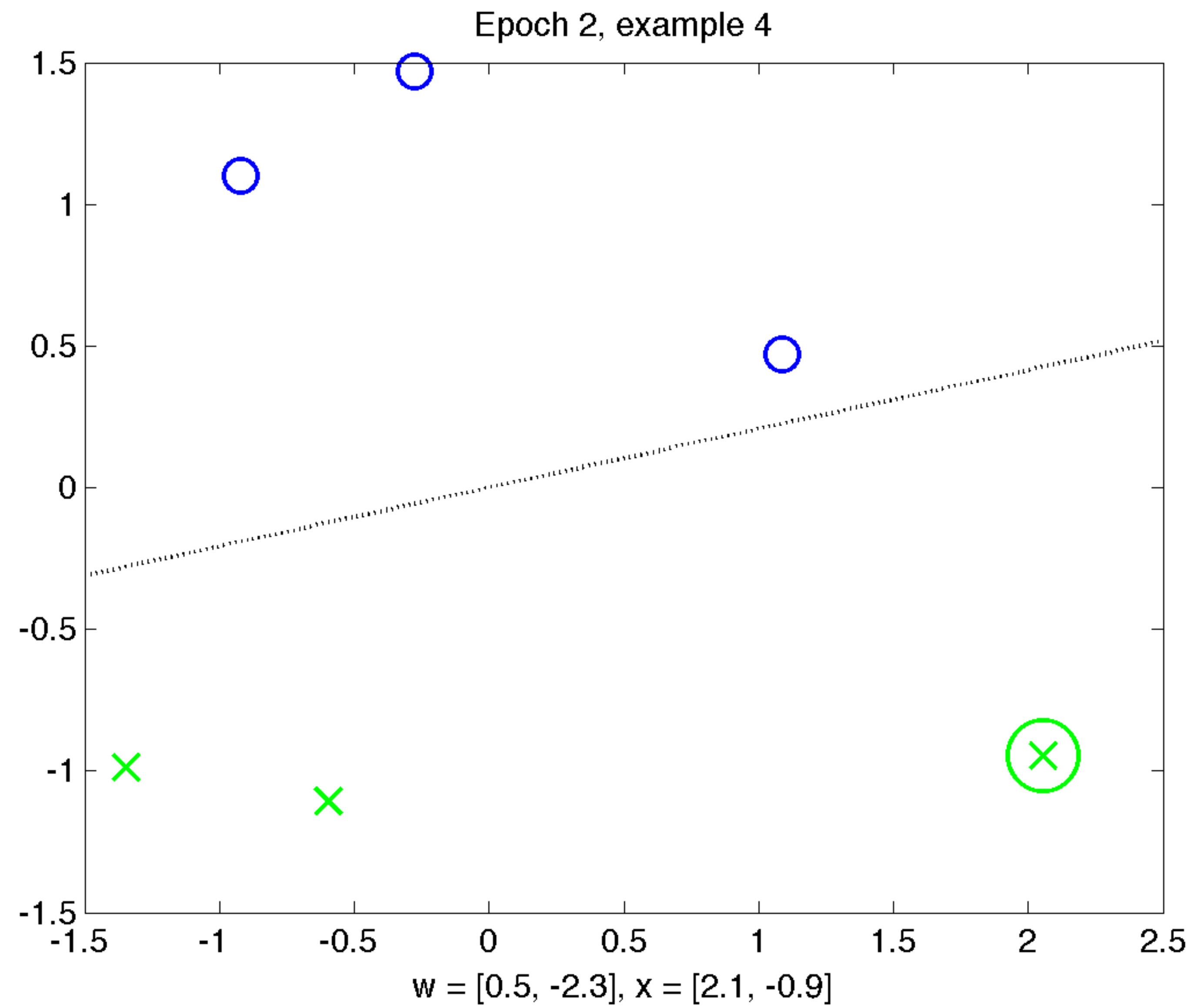


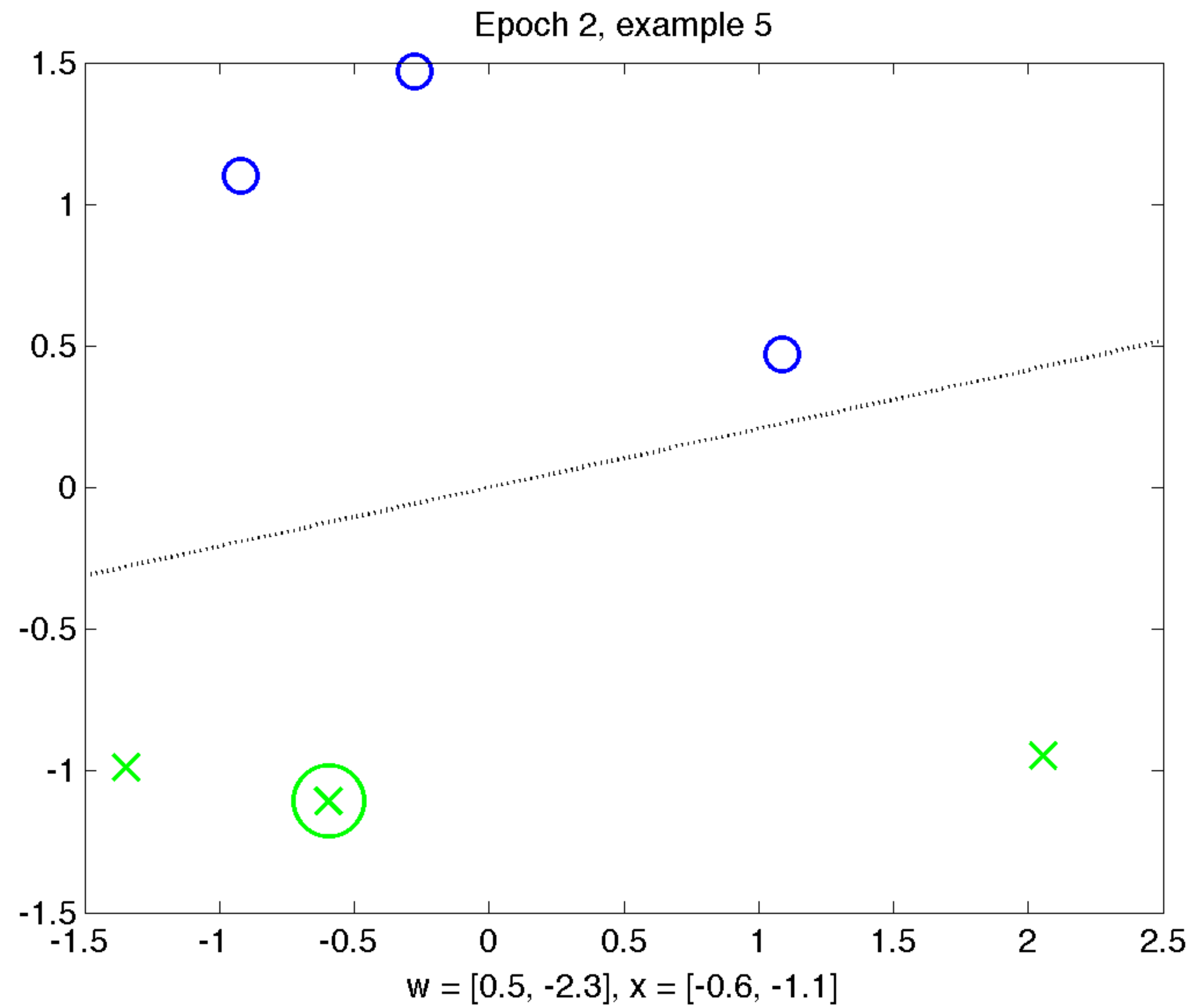


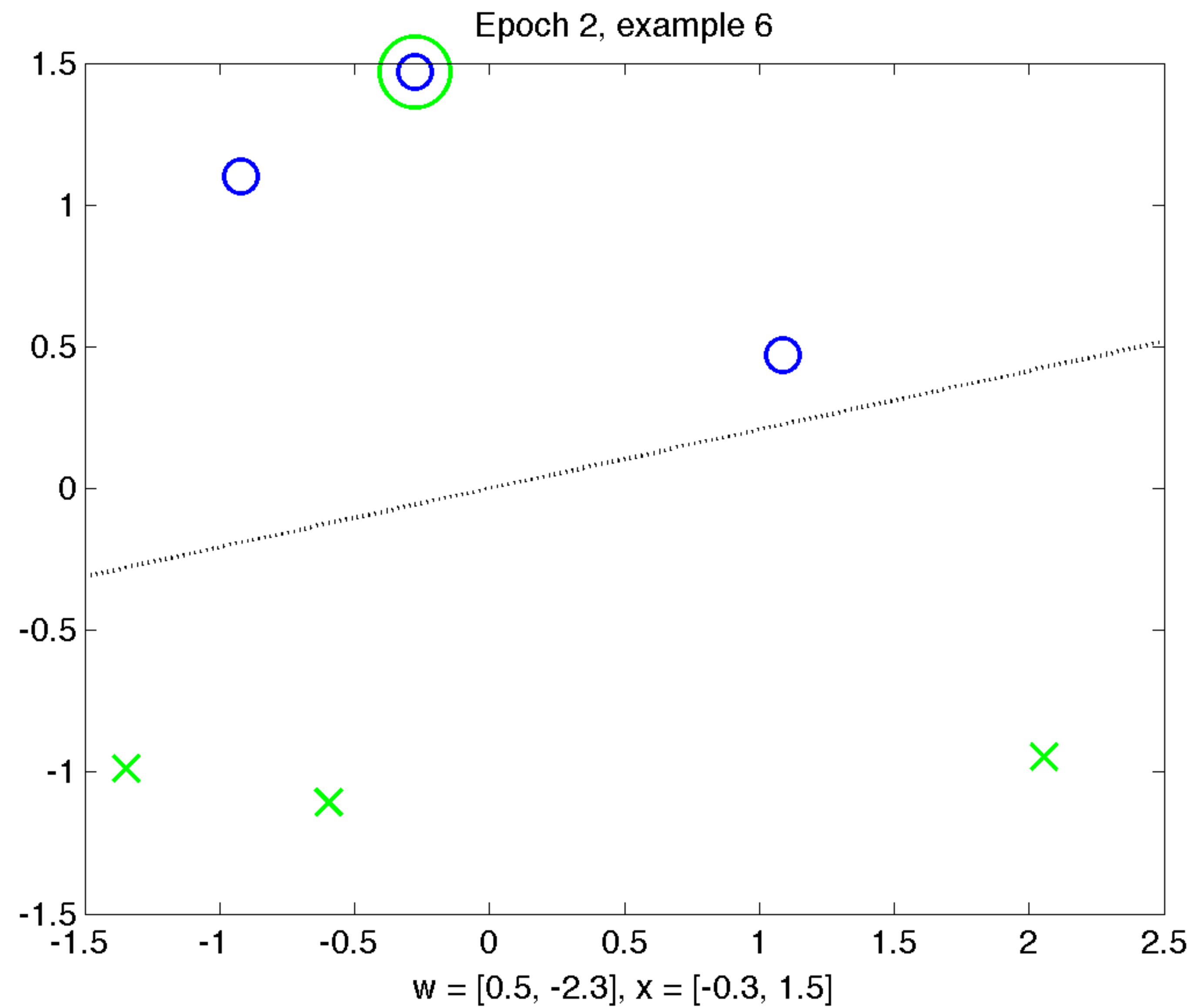












Perceptron Learning

$$\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x})$$

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha \mathbf{y} \mathbf{x} \text{ if } \mathbf{f}(\mathbf{x}) \neq \mathbf{y}$$

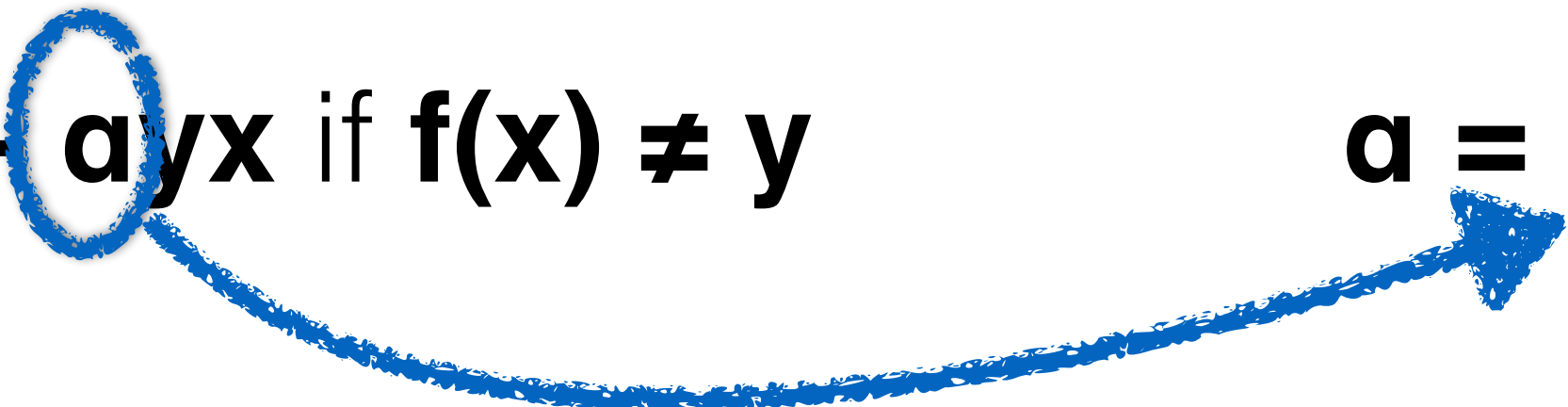
α = learning rate

Perceptron Learning

$$f(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x})$$

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha y \mathbf{x} \text{ if } f(\mathbf{x}) \neq y$$

α = learning rate



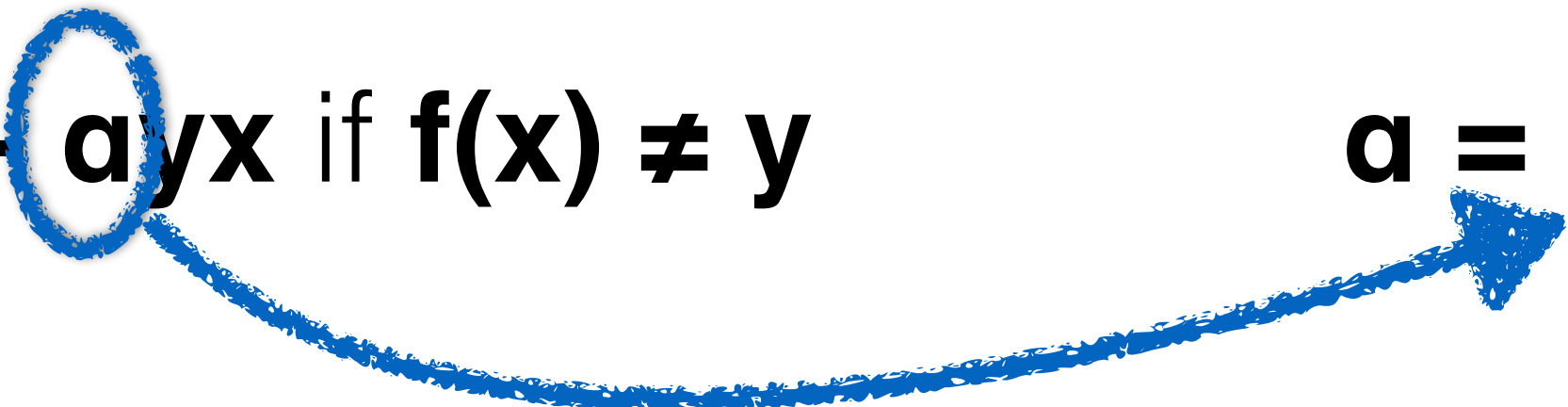
Perceptron Learning

$$\text{sign}(\mathbf{w} \cdot \mathbf{x} + b)$$

$$\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x})$$

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha y \mathbf{x} \text{ if } \mathbf{f}(\mathbf{x}) \neq y$$

α = learning rate



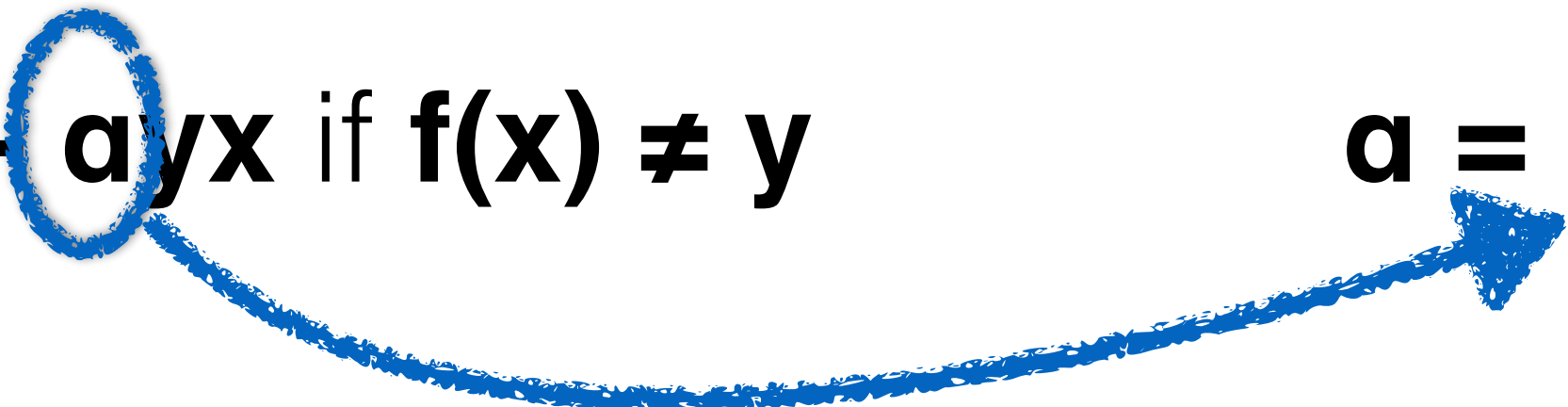
Perceptron Learning

$$\text{sign}(\mathbf{w} \cdot \mathbf{x} + b)$$

$$\mathbf{f}(\mathbf{x}) = \text{sign}(\mathbf{w} \cdot \mathbf{x})$$

$$\mathbf{w} \leftarrow \mathbf{w} + \alpha y \mathbf{x} \text{ if } \mathbf{f}(\mathbf{x}) \neq y$$

α = learning rate

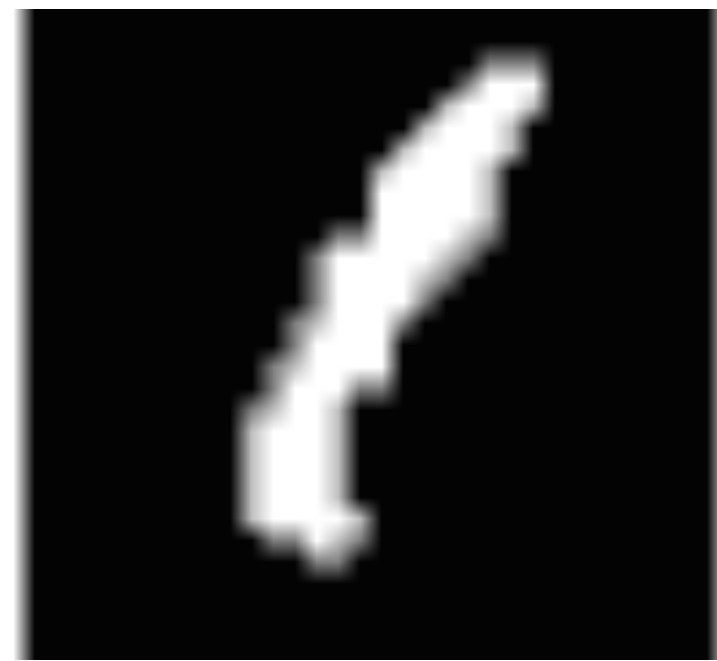


$$b \leftarrow b + \alpha y \text{ if } \mathbf{f}(\mathbf{x}) \neq y$$

Multiclass Output



-1



+1

Multiclass Output



Multiclass Perceptron

Multiclass Perceptron

- Separate weights for each class: $\mathbf{w}^0, \mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3, \mathbf{w}^4, \mathbf{w}^5, \mathbf{w}^6, \dots$

Multiclass Perceptron

- Separate weights for each class: $\mathbf{w}^0, \mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3, \mathbf{w}^4, \mathbf{w}^5, \mathbf{w}^6, \dots$
- Predict according to maximum scoring class:

$$\mathbf{f}(\mathbf{x}) = \operatorname{argmax}_{y'} \mathbf{w}^{y'} \cdot \mathbf{x}$$

Multiclass Perceptron

- Separate weights for each class: $\mathbf{w}^0, \mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3, \mathbf{w}^4, \mathbf{w}^5, \mathbf{w}^6, \dots$
- Predict according to maximum scoring class:

$$\mathbf{f}(\mathbf{x}) = \operatorname{argmax}_{y'} \mathbf{w}^{y'} \cdot \mathbf{x}$$

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{\mathbf{f}(\mathbf{x})} \leftarrow \mathbf{w}^{\mathbf{f}(\mathbf{x})} - \mathbf{x}$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

x



Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

x



y = 2

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

x



$$\mathbf{y} = 2$$

$$\mathbf{f}(\mathbf{x}) = 6$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 6$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x} \quad (\text{make } \mathbf{w}^2 \cdot \mathbf{x} \text{ greater})$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 6$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x} \quad (\text{make } \mathbf{w}^2 \cdot \mathbf{x} \text{ greater})$$

$$\mathbf{w}^6 \leftarrow \mathbf{w}^6 - \mathbf{x} \quad (\text{make } \mathbf{w}^6 \cdot \mathbf{x} \text{ lesser})$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

x



y = 2

Multiclass Perceptron

- Learning update:

$$\mathbf{w}_y \leftarrow \mathbf{w}_y + \mathbf{x}$$

$$\mathbf{w}_{f(\mathbf{x})} \leftarrow \mathbf{w}_{f(\mathbf{x})} - \mathbf{x}$$

x



$$\mathbf{y} = 2$$

$$\mathbf{f}(\mathbf{x}) = 2$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 2$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x}$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 2$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x}$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 - \mathbf{x}$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 2$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x} - \mathbf{x}$$

Multiclass Perceptron

- Learning update:

$$\mathbf{w}^y \leftarrow \mathbf{w}^y + \mathbf{x}$$

$$\mathbf{w}^{f(\mathbf{x})} \leftarrow \mathbf{w}^{f(\mathbf{x})} - \mathbf{x}$$

$\mathbf{x}$



$$y = 2$$

$$f(\mathbf{x}) = 2$$

$$\mathbf{w}^2 \leftarrow \mathbf{w}^2 + \mathbf{x} - \mathbf{x} = \mathbf{w}^2$$

Perceptron Summary

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**
- Learn by updating when wrong

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**
- Learn by updating when wrong
 - If score too low, make score higher

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**
- Learn by updating when wrong
 - If score too low, make score higher
 - If score too high, make score lower

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**
- Learn by updating when wrong
 - If score too low, make score higher
 - If score too high, make score lower
- Multiclass: use **multiple weight vectors**, choose max scoring

Perceptron Summary

- Linear classifier multiplies **weights** by input **features**
- Learn by updating when wrong
 - If score too low, make score higher
 - If score too high, make score lower
- Multiclass: use **multiple weight vectors**, choose max scoring
 - Learn by adjusting score of predicted class and true class