Practice-1

Machine Learning Techniques

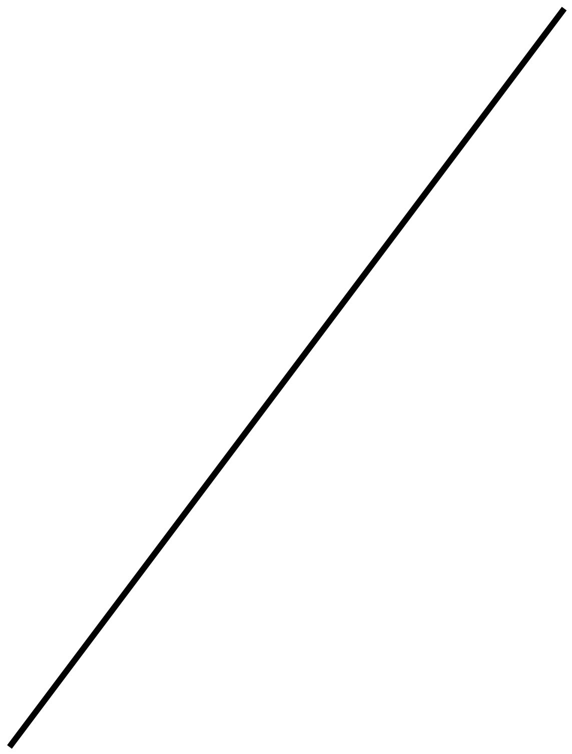
Karthik Thiagarajan

Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

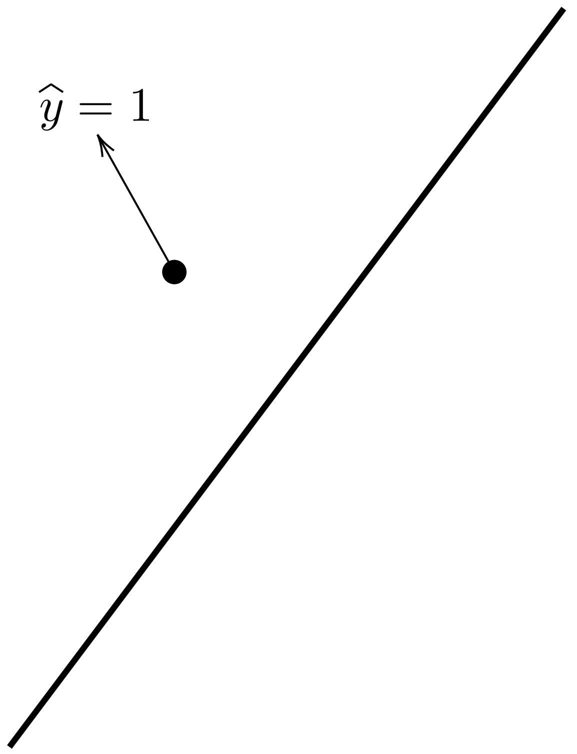
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary



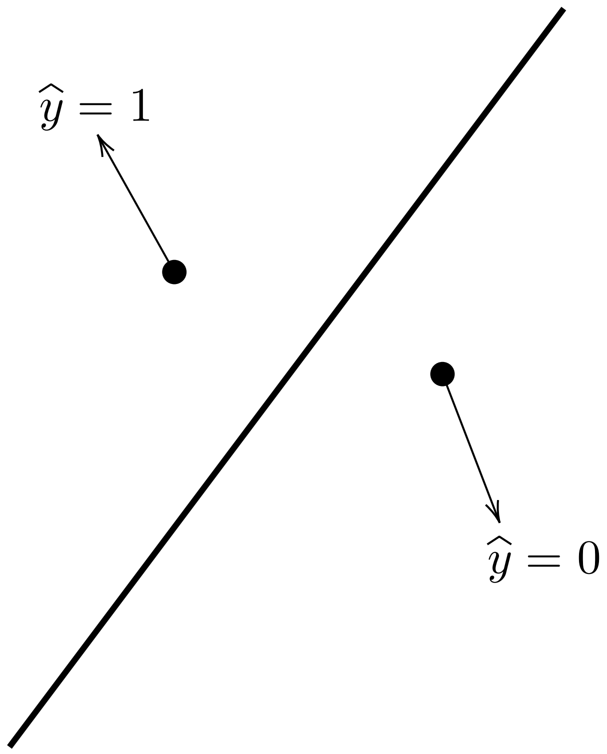
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary



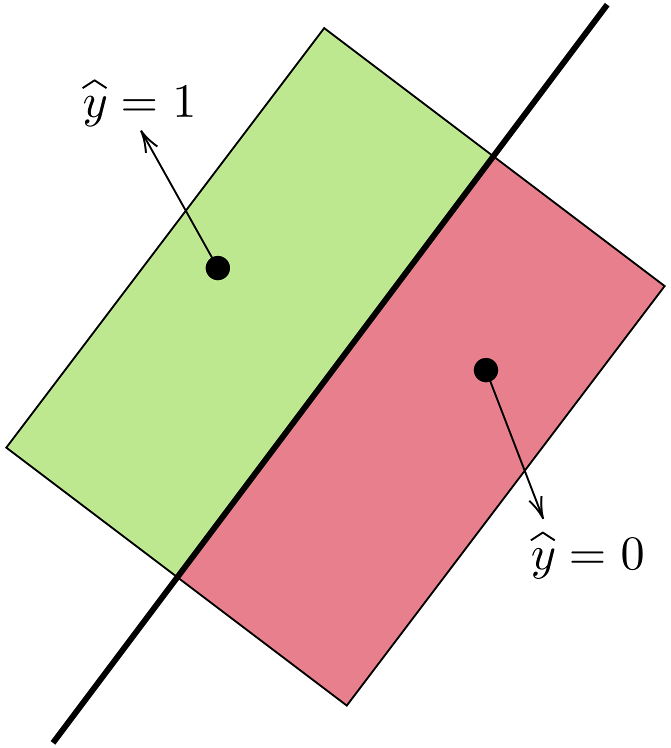
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary



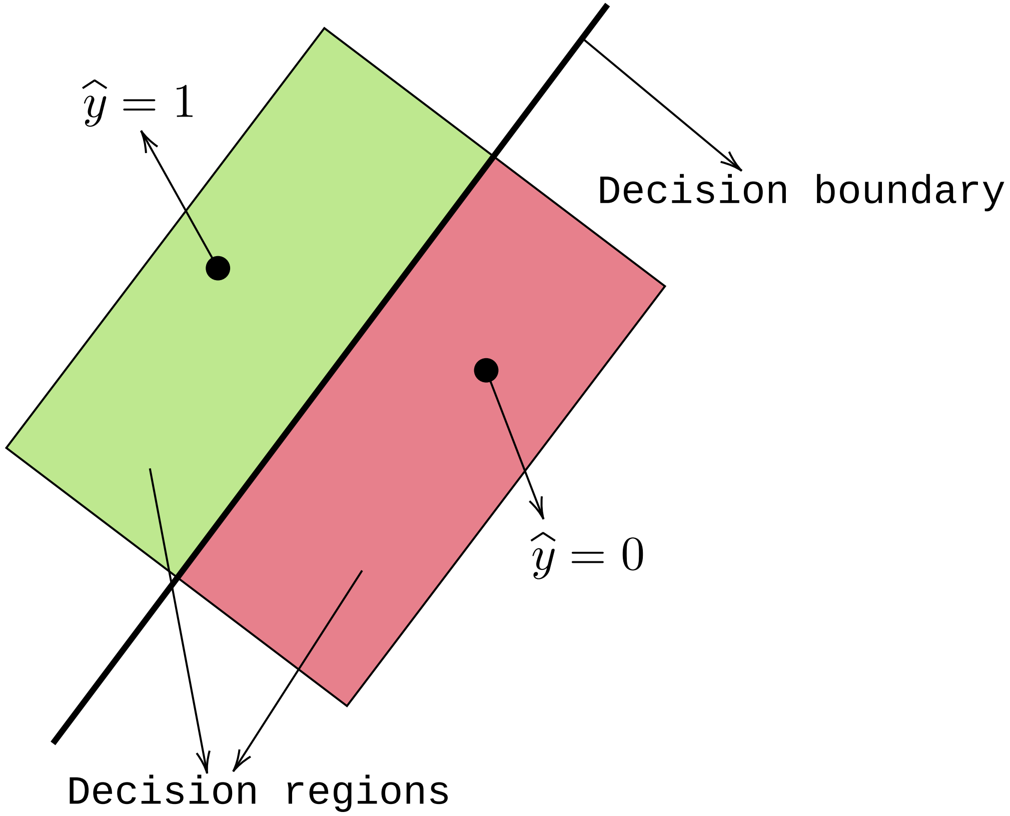
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary



Q-1

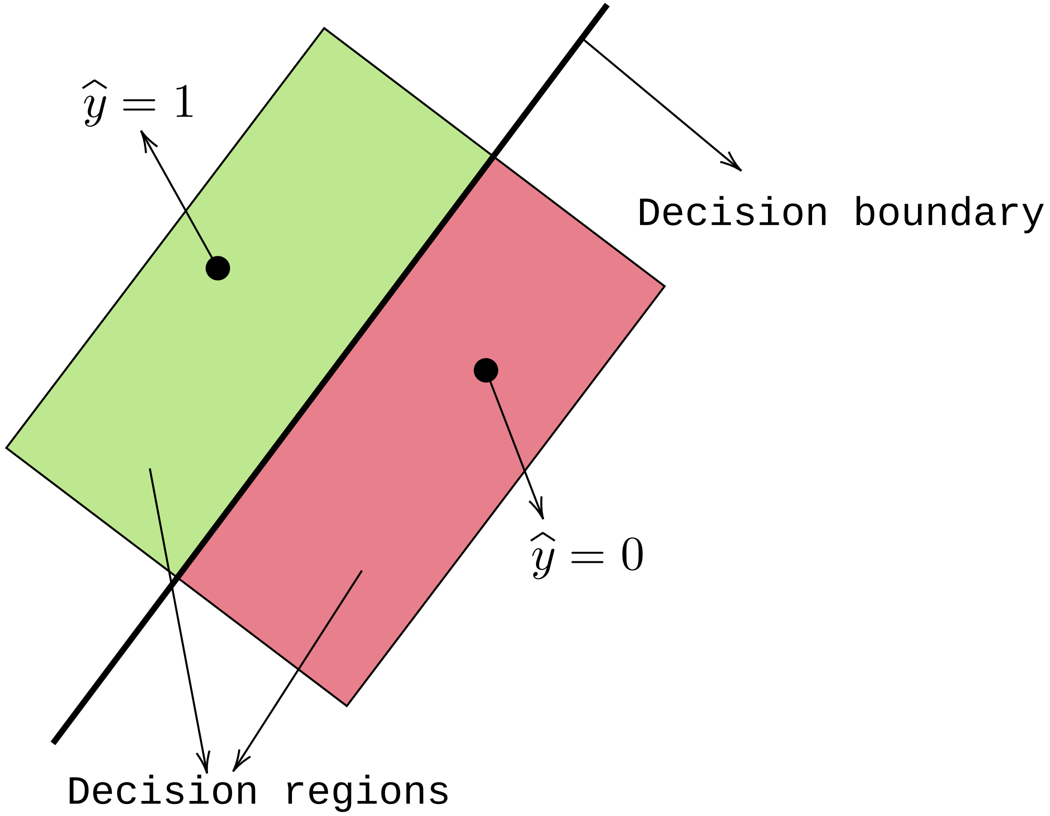
- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary



Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

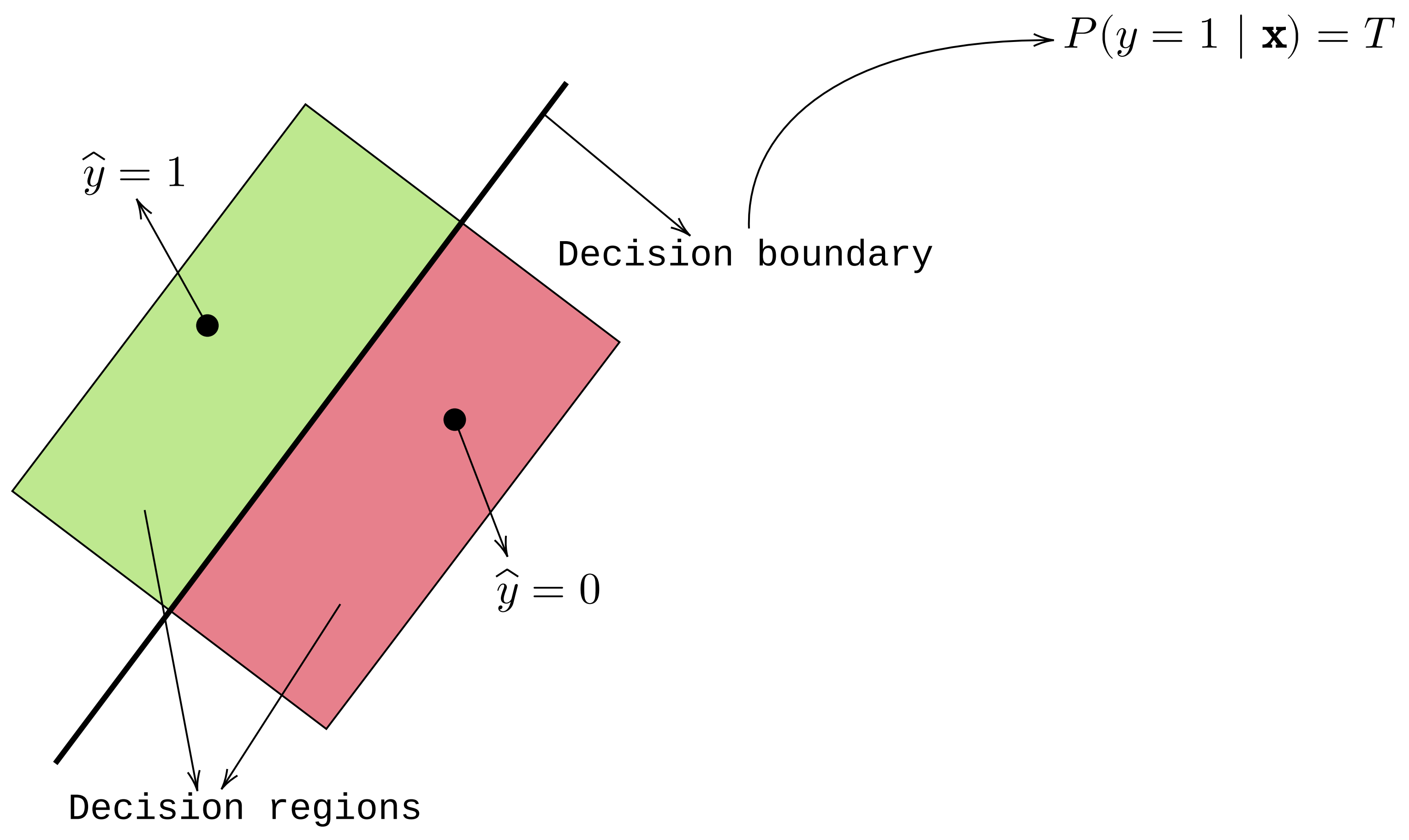
$$\hat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$



Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\hat{y} = \begin{cases} 1, & P(y = 1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$



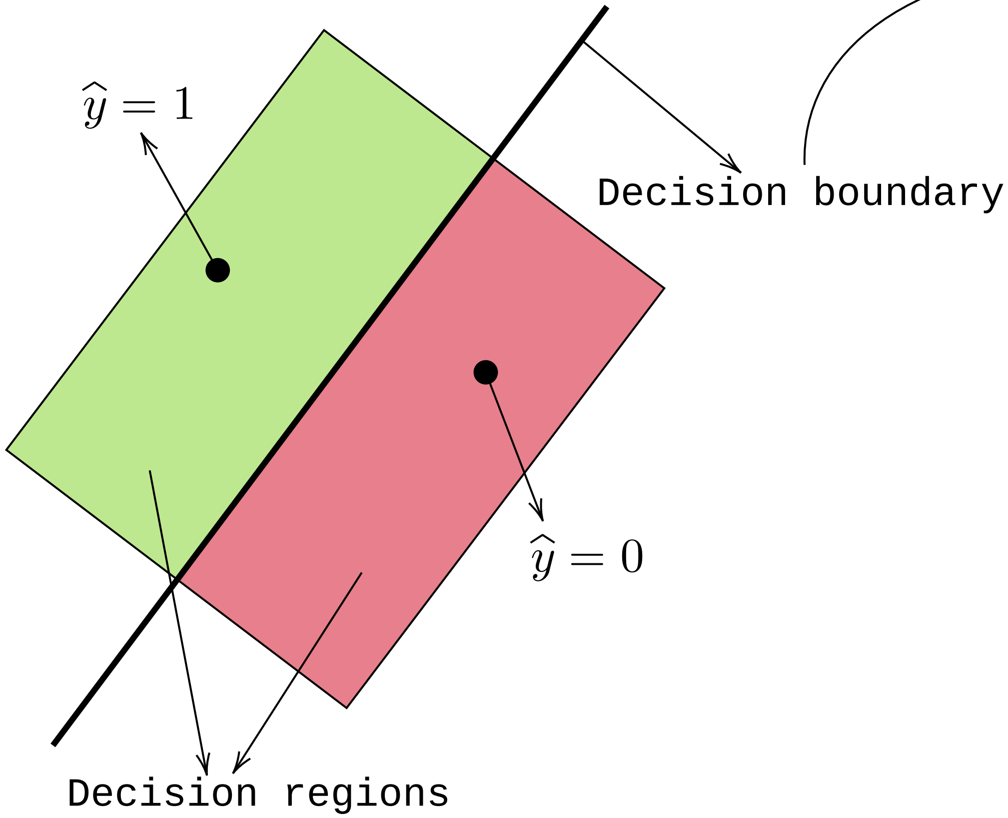
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\hat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$

$P(y=1 \mid \mathbf{x}) = T$

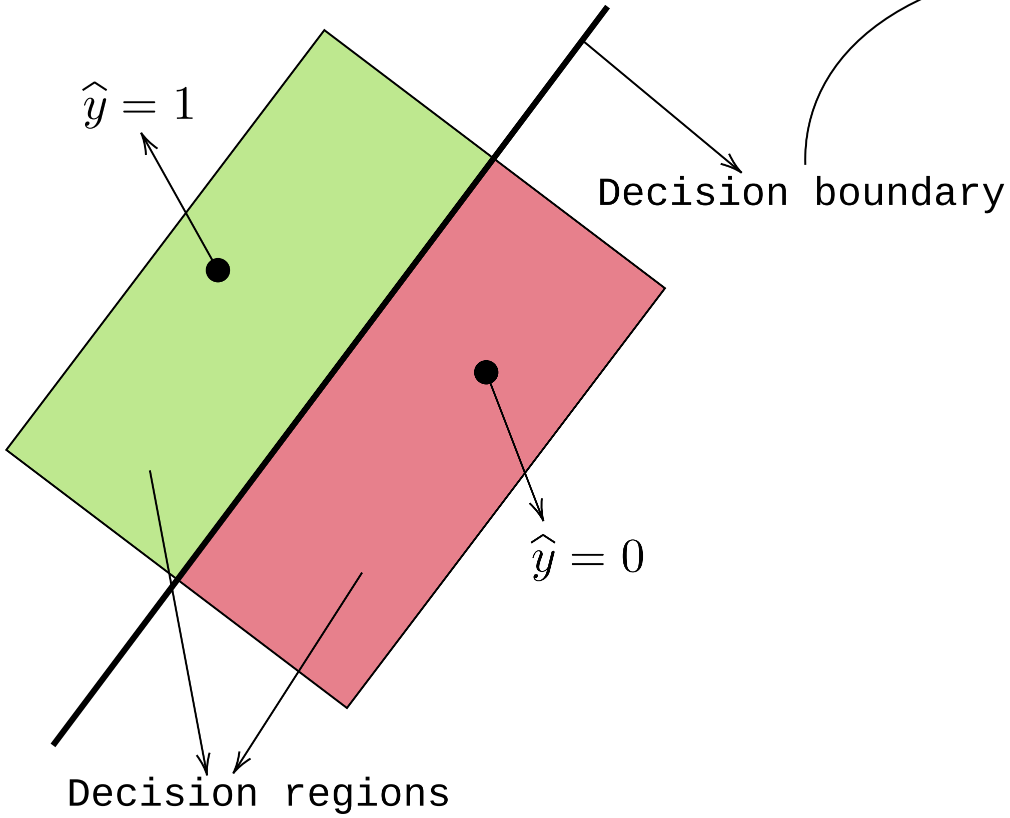
$\sigma(\mathbf{w}^T\mathbf{x}) = T$



Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$



$$P(y=1 \mid \mathbf{x}) = T$$

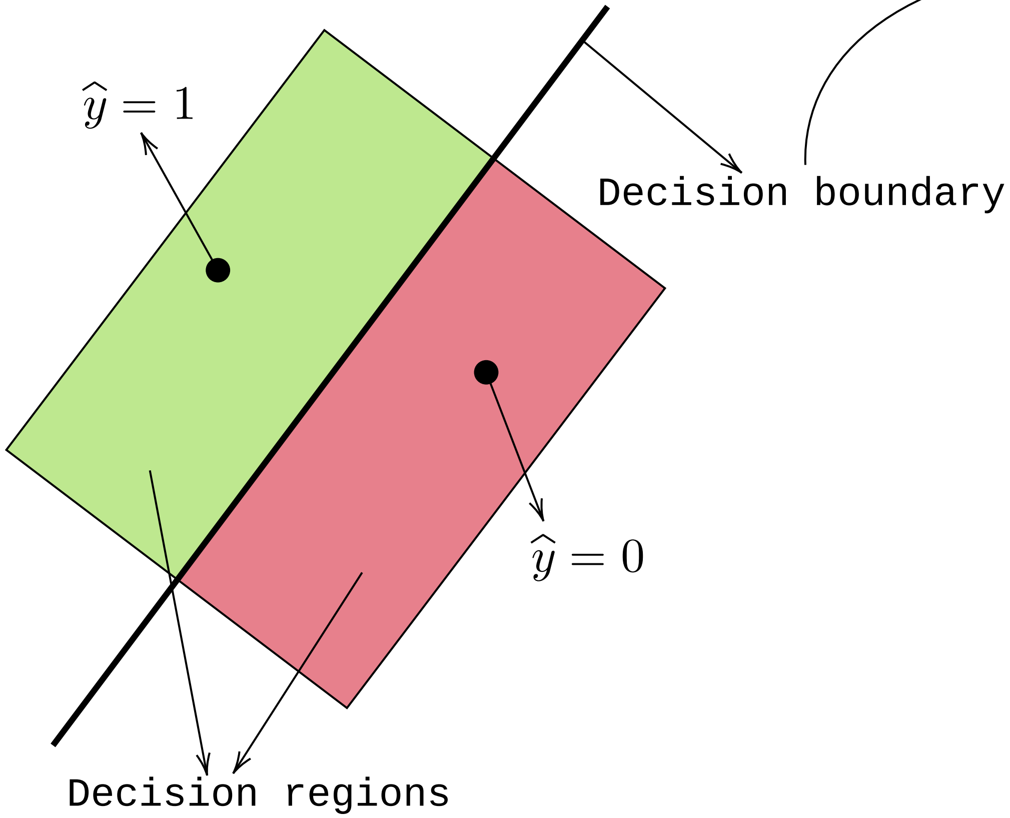
$$\sigma(\mathbf{w}^T\mathbf{x}) = T$$

$$\frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} = T$$

Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \texttt{otherwise} \end{cases}$$



$$P(y=1 \mid \mathbf{x}) = T$$

$$\sigma(\mathbf{w}^T\mathbf{x}) = T$$

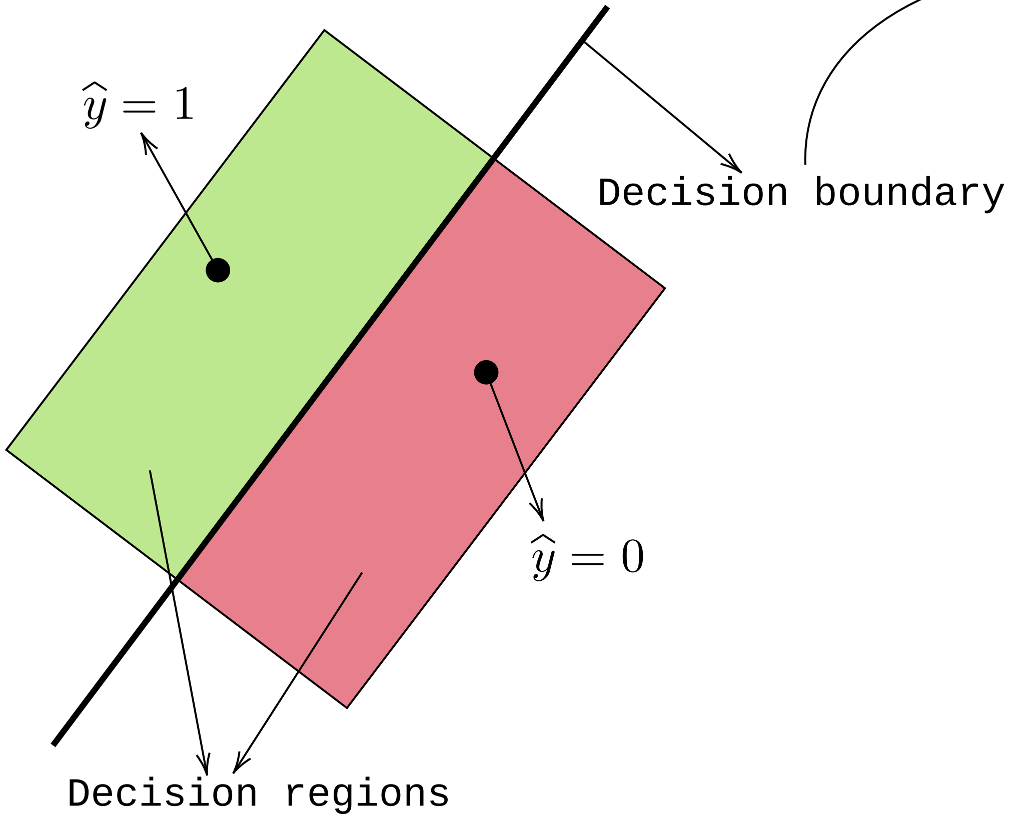
$$\frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} = T$$

$$1+e^{-\mathbf{w}^T\mathbf{x}} = \frac{1}{T}$$

Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \texttt{otherwise} \end{cases}$$



$$P(y=1 \mid \mathbf{x}) = T$$

$$\begin{aligned} \sigma\left(\mathbf{w}^T\mathbf{x}\right) &= T \\ \frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} &= T \\ 1+e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\ e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \end{aligned}$$

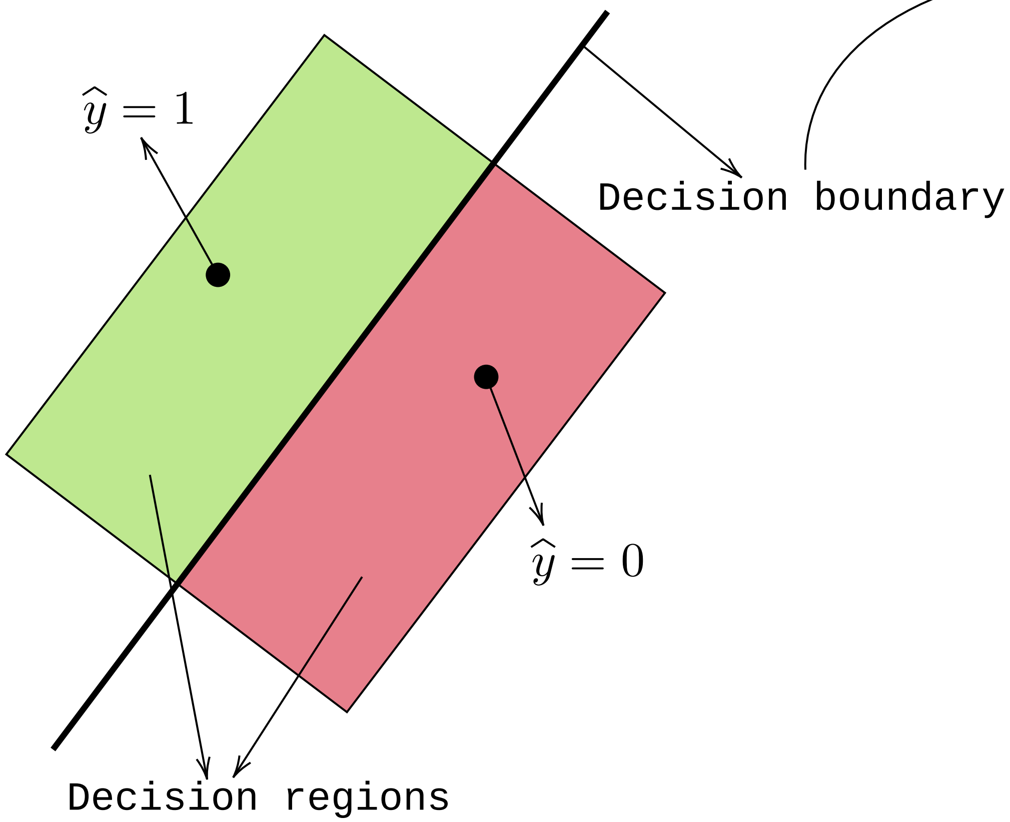
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \texttt{otherwise} \end{cases}$$

$$P(y=1 \mid \mathbf{x}) = T$$

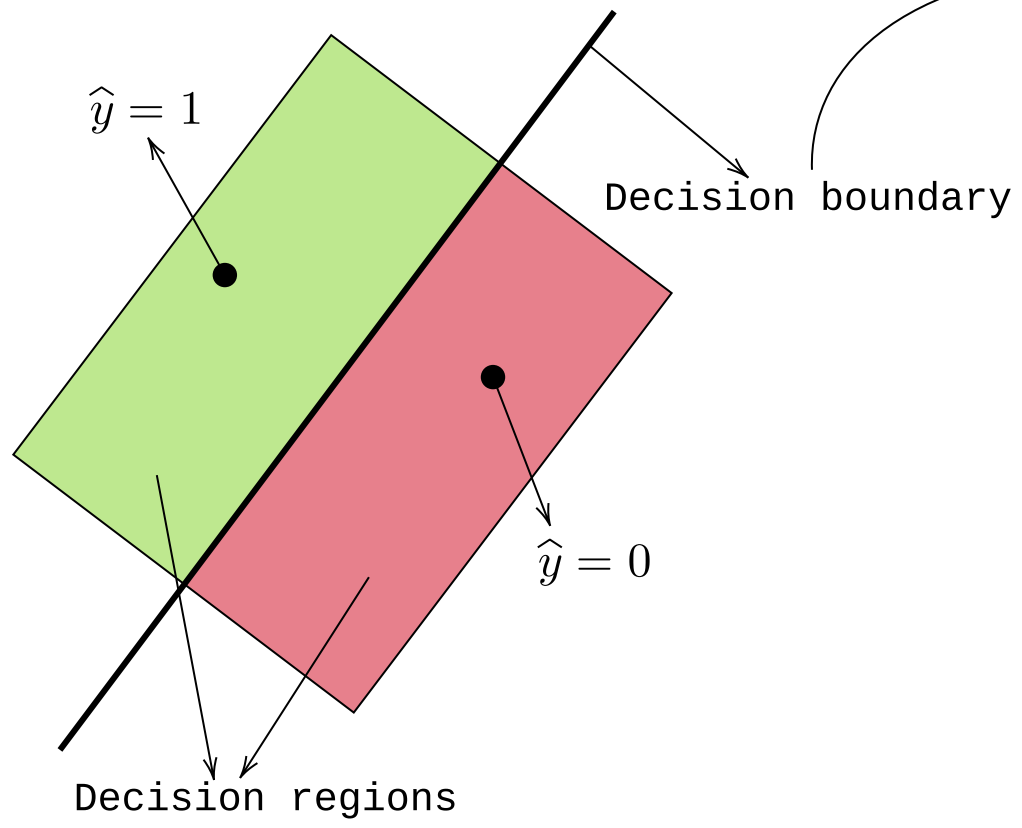
$$\begin{aligned} \sigma(\mathbf{w}^T\mathbf{x}) &= T \\ \frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} &= T \\ 1+e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\ e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \\ -\mathbf{w}^T\mathbf{x} &= \ln\left(\frac{1}{T} - 1\right) \end{aligned}$$



Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\hat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$



$$P(y=1 \mid \mathbf{x}) = T$$

$$\begin{aligned}
\sigma(\mathbf{w}^T\mathbf{x}) &= T \\
\frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} &= T \\
1+e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\
e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \\
-\mathbf{w}^T\mathbf{x} &= \ln\left(\frac{1}{T} - 1\right) \\
\mathbf{w}^T\mathbf{x} &= -\ln\left(\frac{1}{T} - 1\right)
\end{aligned}$$

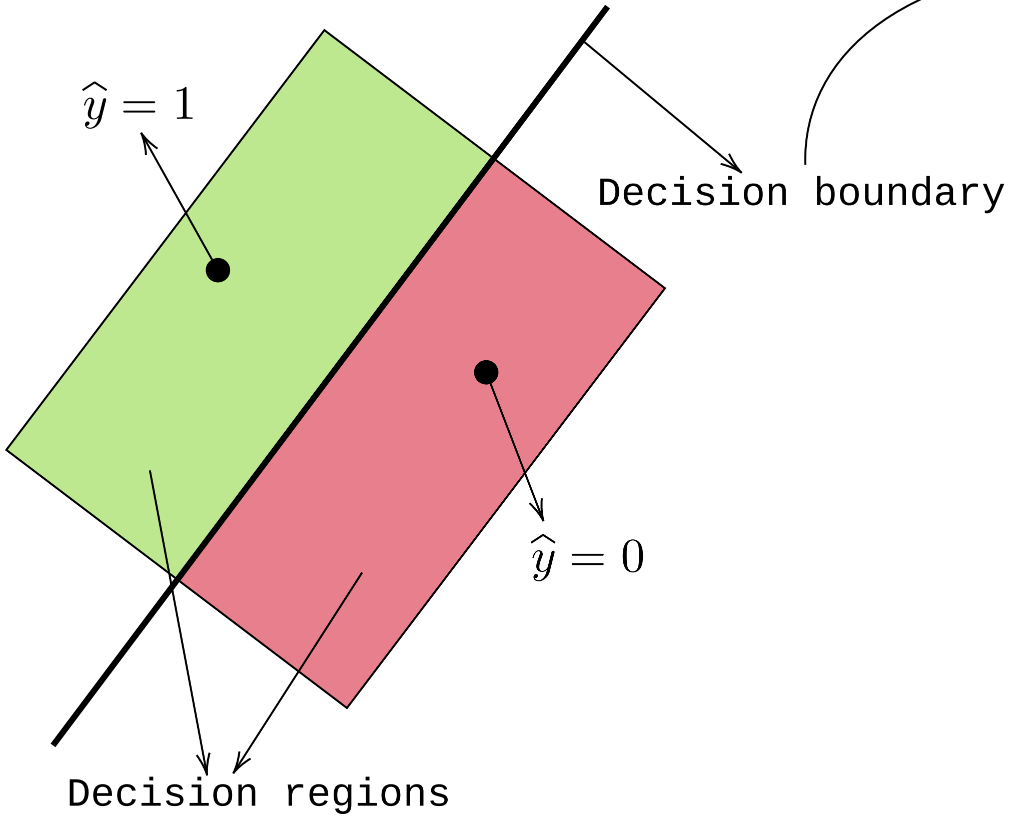
Q-1

- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

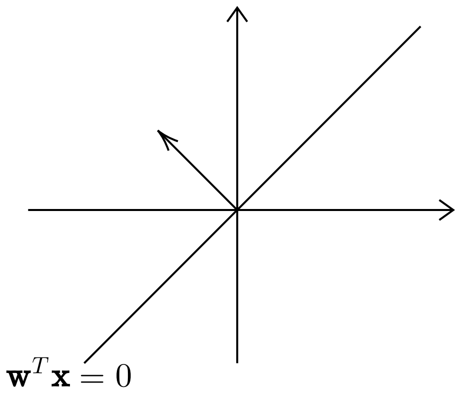
$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$

$$P(y=1 \mid \mathbf{x}) = T$$

$$\begin{aligned}
\sigma(\mathbf{w}^T\mathbf{x}) &= T \\
\frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} &= T \\
1+e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\
e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \\
-\mathbf{w}^T\mathbf{x} &= \ln\left(\frac{1}{T} - 1\right) \\
\mathbf{w}^T\mathbf{x} &= -\ln\left(\frac{1}{T} - 1\right)
\end{aligned}$$



$T = 0.5$



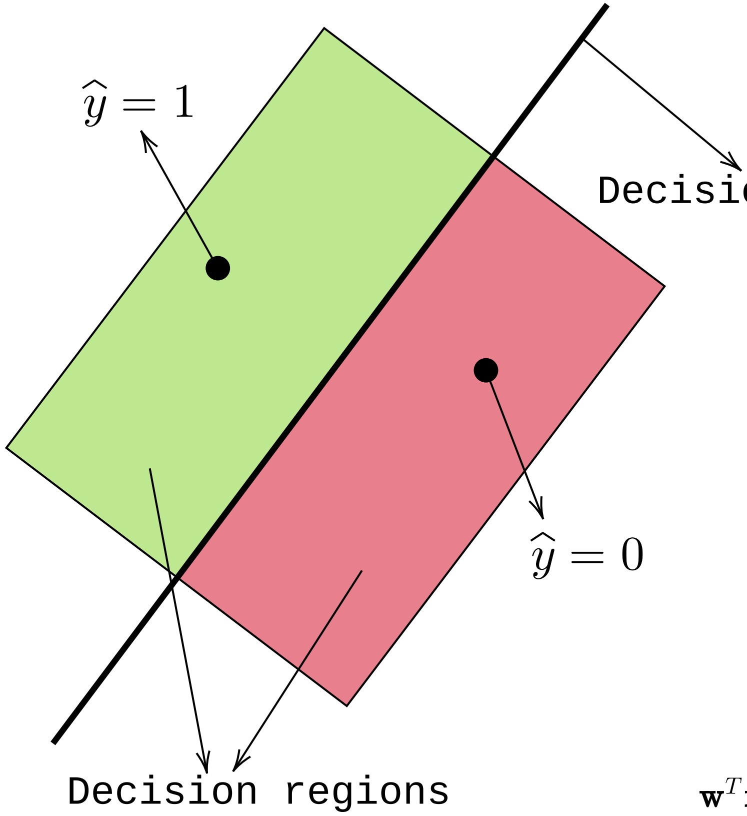
Q-1

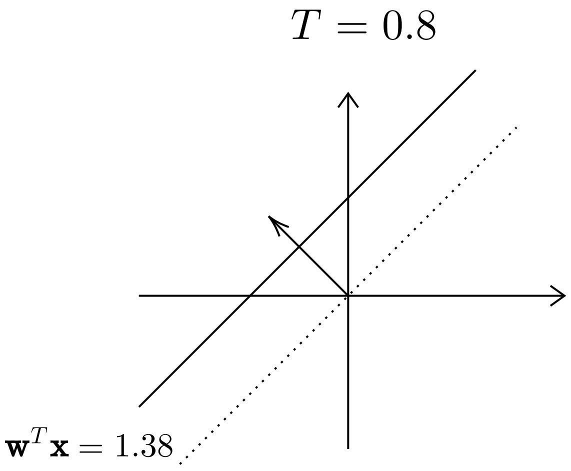
- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

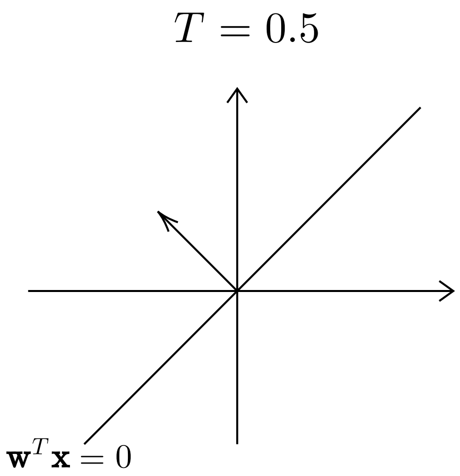
$$\widehat{y} = \begin{cases} 1, & P(y = 1 \mid \mathbf{x}) \geqslant T \\ 0, & \texttt{otherwise} \end{cases}$$

$$P(y = 1 \mid \mathbf{x}) = T$$

$$\begin{aligned} \sigma\left(\mathbf{w}^T\mathbf{x}\right) &= T \\ \frac{1}{1 + e^{-\mathbf{w}^T\mathbf{x}}} &= T \\ 1 + e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\ e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \\ -\mathbf{w}^T\mathbf{x} &= \ln\left(\frac{1}{T} - 1\right) \\ \mathbf{w}^T\mathbf{x} &= -\ln\left(\frac{1}{T} - 1\right) \end{aligned}$$



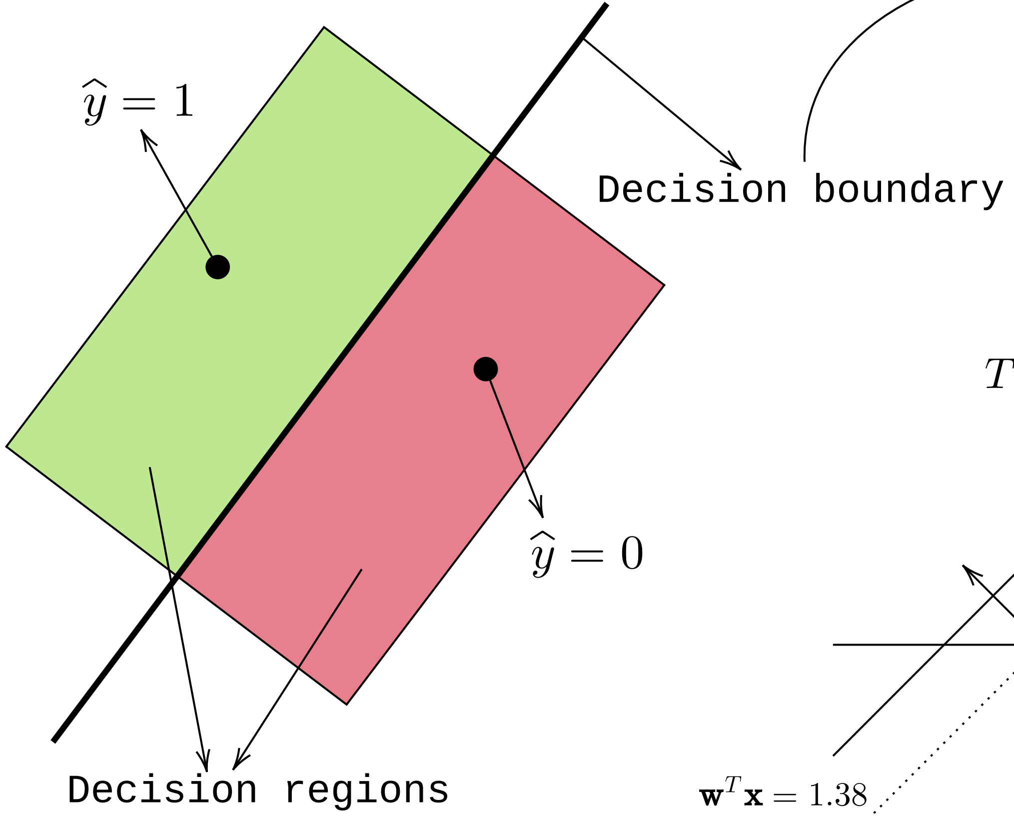




Q-1

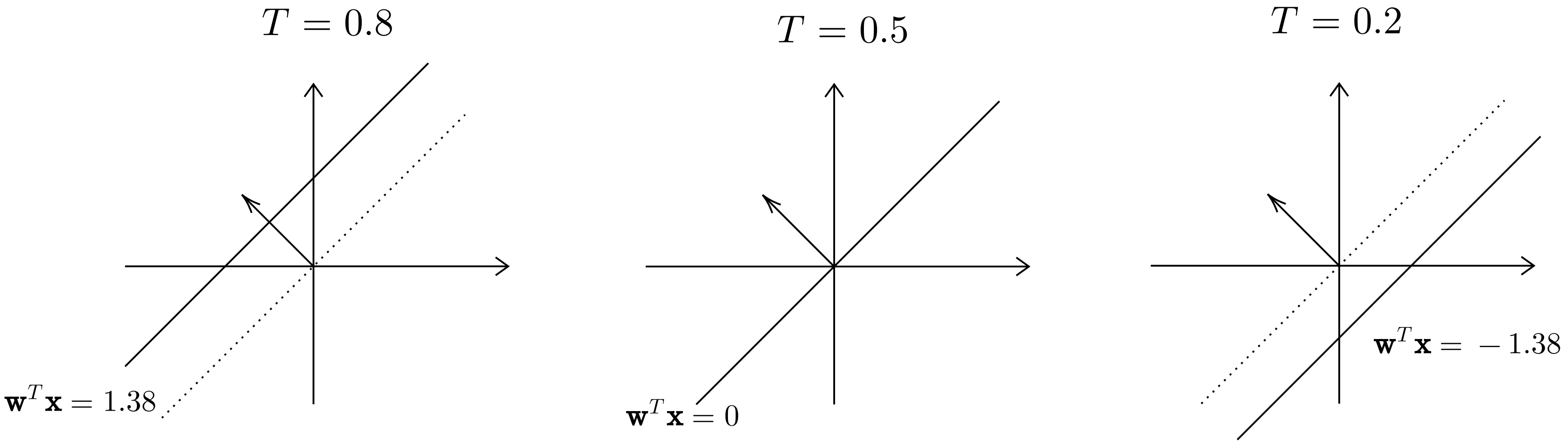
- Logistic Regression
- Threshold for prediction is T
- Find the decision boundary

$$\widehat{y} = \begin{cases} 1, & P(y=1 \mid \mathbf{x}) \geqslant T \\ 0, & \text{otherwise} \end{cases}$$



$$P(y=1 \mid \mathbf{x}) = T$$

$$\begin{aligned} \sigma(\mathbf{w}^T\mathbf{x}) &= T \\ \frac{1}{1+e^{-\mathbf{w}^T\mathbf{x}}} &= T \\ 1+e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} \\ e^{-\mathbf{w}^T\mathbf{x}} &= \frac{1}{T} - 1 \\ -\mathbf{w}^T\mathbf{x} &= \ln\left(\frac{1}{T} - 1\right) \\ \mathbf{w}^T\mathbf{x} &= -\ln\left(\frac{1}{T} - 1\right) \end{aligned}$$



Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \quad 1]^T$
- Hinge loss

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\left[1 - \left(\mathbf{w}^T \mathbf{x}_i\right) y_i, 0\right]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \left(\mathbf{w}^T \mathbf{x}\right) y$	L
2	1	1		
-2	1	1		
-1	2	1		
0	2	-1		
1	-1	-1		
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \ \ 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \big(\mathbf{w}^T \mathbf{x}\big) y$	L
2	1	1	0	0
-2	1	1		
-1	2	1		
0	2	-1		
1	-1	-1		
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \ \ 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \big(\mathbf{w}^T \mathbf{x}\big) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1		
0	2	-1		
1	-1	-1		
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \ \ 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \big(\mathbf{w}^T \mathbf{x}\big) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1		
1	-1	-1		
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\left[1 - \left(\mathbf{w}^T \mathbf{x}_i\right) y_i, 0\right]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \left(\mathbf{w}^T \mathbf{x}\right) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1	3	3
1	-1	-1		
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \big(\mathbf{w}^T \mathbf{x}\big) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1	3	3
1	-1	-1	0	0
2	-2	-1		
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\Big[1 - \big(\mathbf{w}^T \mathbf{x}_i\big) y_i, 0\Big]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \big(\mathbf{w}^T \mathbf{x}\big) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1	3	3
1	-1	-1	0	0
2	-2	-1	-1	0
-2	0	-1		

Q-2

- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\left[1 - \left(\mathbf{w}^T \mathbf{x}_i\right) y_i, 0\right]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \left(\mathbf{w}^T \mathbf{x}\right) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1	3	3
1	-1	-1	0	0
2	-2	-1	-1	0
-2	0	-1	1	1

Q-2

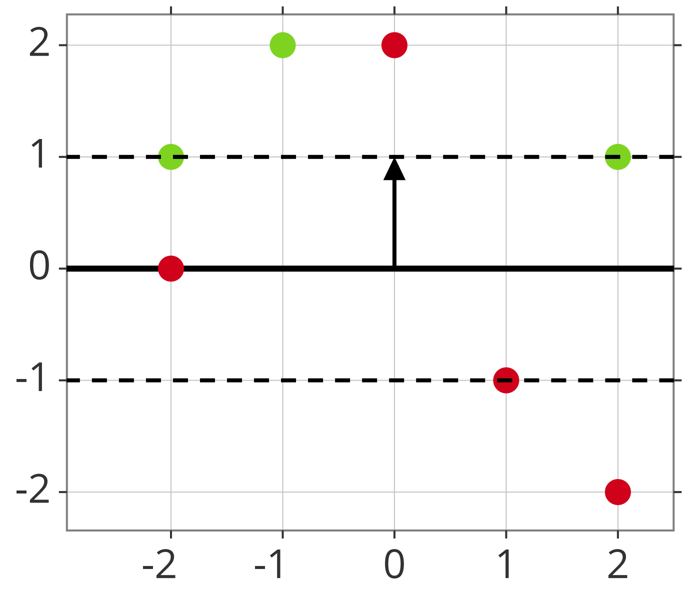
- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \quad 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\left[1 - \left(\mathbf{w}^T \mathbf{x}_i\right) y_i, 0\right]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
-2	1	1
-1	2	1
0	2	-1
1	-1	-1
2	-2	-1
-2	0	-1

x_1	x_2	y	$1 - \left(\mathbf{w}^T \mathbf{x}\right) y$	L
2	1	1	0	0
-2	1	1	0	0
-1	2	1	-1	0
0	2	-1	3	3
1	-1	-1	0	0
2	-2	-1	-1	0
-2	0	-1	1	1



Q-2

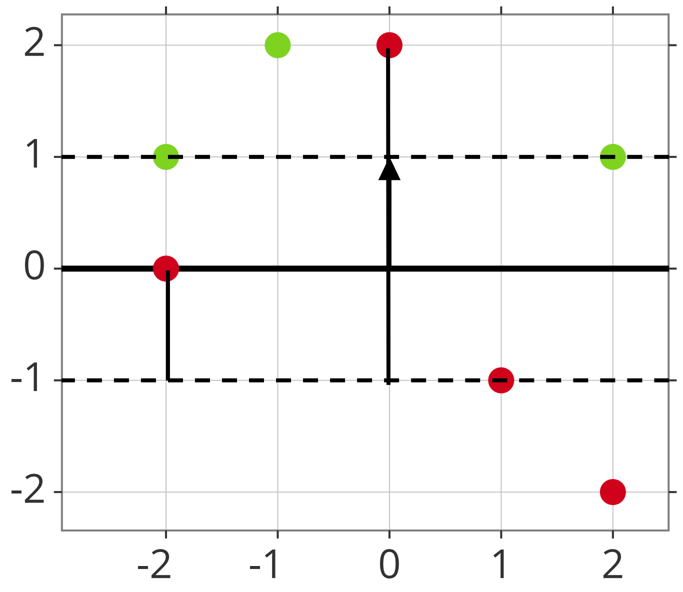
- Soft-margin, Linear-SVM
- $\mathbf{w} = [0 \;\; 1]^T$
- Hinge loss

$$L(\mathbf{X}, \mathbf{y}, \mathbf{w}) = \sum_{i=1}^{n} \max\left[1 - \left(\mathbf{w}^T \mathbf{x}_i\right) y_i, 0\right]$$

$\mathbf{X}, \mathbf{y} =$

x_1	x_2	y
2	1	1
−2	1	1
−1	2	1
0	2	−1
1	−1	−1
2	−2	−1
−2	0	−1

x_1	x_2	y	$1 - \left(\mathbf{w}^T \mathbf{x}\right) y$	L
2	1	1	0	0
−2	1	1	0	0
−1	2	1	−1	0
0	2	−1	3	3
1	−1	−1	0	0
2	−2	−1	−1	0
−2	0	−1	1	1



Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T \right)^T$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$|\mathbf{C} - \lambda \mathbf{I}|$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$|\mathbf{C} - \lambda \mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix}$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$|\mathbf{C} - \lambda \mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix}$$

$$= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda)$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$|\mathbf{C} - \lambda \mathbf{I}| = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix}$$

$$= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda)$$

$$= -(1-\lambda)[\lambda(1-\lambda) + 1]$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} |\mathbf{C} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} \\ &= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda) \\ &= -(1-\lambda)[\lambda(1-\lambda)+1] \\ &= (1-\lambda)\left(\lambda^2 - \lambda - 1\right) \end{aligned}$$

Q-3

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} |\mathbf{C} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} \\ &= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda) \\ &= -(1-\lambda)[\lambda(1-\lambda)+1] \\ &= (1-\lambda)\left(\lambda^2 - \lambda - 1\right) \end{aligned}$$

$$\lambda_1 = \frac{1+\sqrt{5}}{2},$$

$$\lambda_2 = 1$$

$$\lambda_3 = \frac{1-\sqrt{5}}{2}$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} |\mathbf{C} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} \\ &= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda) \\ &= -(1-\lambda)[\lambda(1-\lambda) + 1] \\ &= (1-\lambda)\big(\lambda^2 - \lambda - 1\big) \end{aligned}$$

$$\lambda_1 = \frac{1+\sqrt{5}}{2},$$

$$\lambda_2 = 1$$

$$\lambda_3 = \frac{1-\sqrt{5}}{2}$$

$$\lambda_3 < 0$$

Q-3 Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 4 & 3 & 0 \\ 1 & 9 & 2 \end{bmatrix}$$

$$\mathbf{C}^T \neq \mathbf{C}$$

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$\mathbf{C}^T = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\mathbf{x}_i \mathbf{x}_i^T\right)^T$$

$$= \frac{1}{n} \cdot \sum_{i=}^{n} \left(\mathbf{x}_i^T\right)^T \mathbf{x}_i^T$$

$$= \frac{1}{n} \cdot \sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T$$

$$= \mathbf{C}$$

Valid covariance matrix?

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} |\mathbf{C} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} \\ &= (1-\lambda)(1-\lambda)(-\lambda) - (1-\lambda) \\ &= -(1-\lambda)[\lambda(1-\lambda)+1] \\ &= (1-\lambda)\left(\lambda^2 - \lambda - 1\right) \end{aligned}$$

$$\lambda_1 = \frac{1+\sqrt{5}}{2},$$

$$\lambda_2 = 1$$

$$\lambda_3 = \frac{1-\sqrt{5}}{2}$$

$$\lambda_3 < 0$$

$$\begin{aligned} \mathbf{A} &= \mathbf{Q}\mathbf{D}\mathbf{Q}^T \\ &= \mathbf{Q}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{Q}^T \\ &= \mathbf{B}\mathbf{B}^T \end{aligned}$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

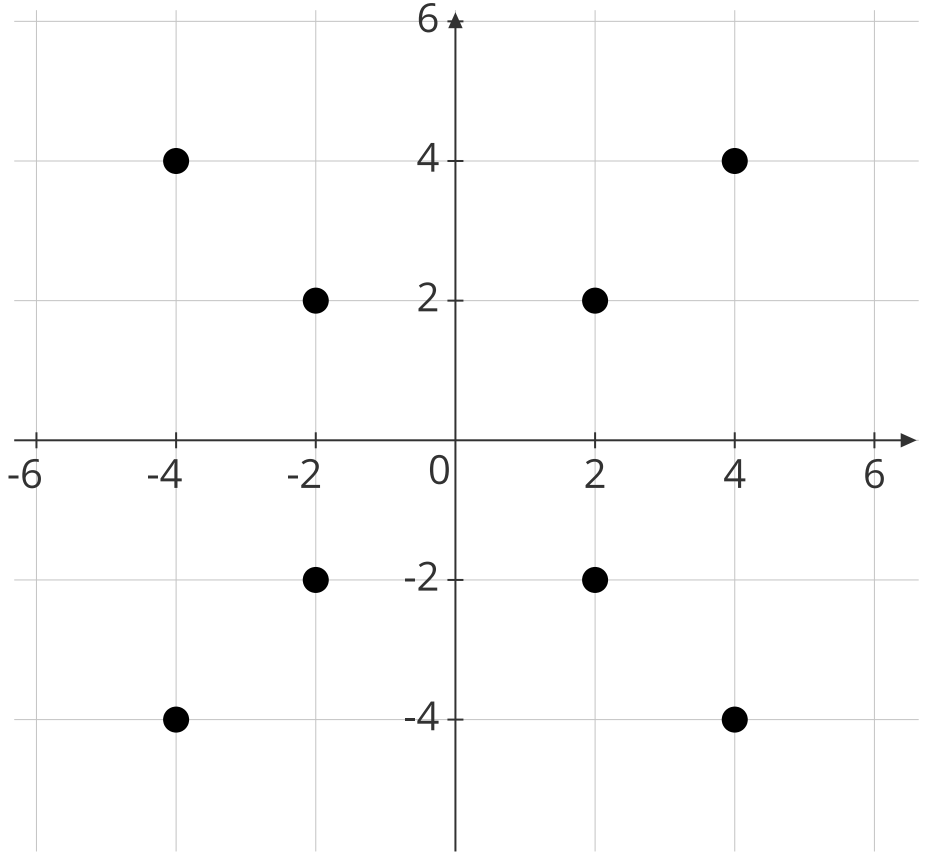
$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$



$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

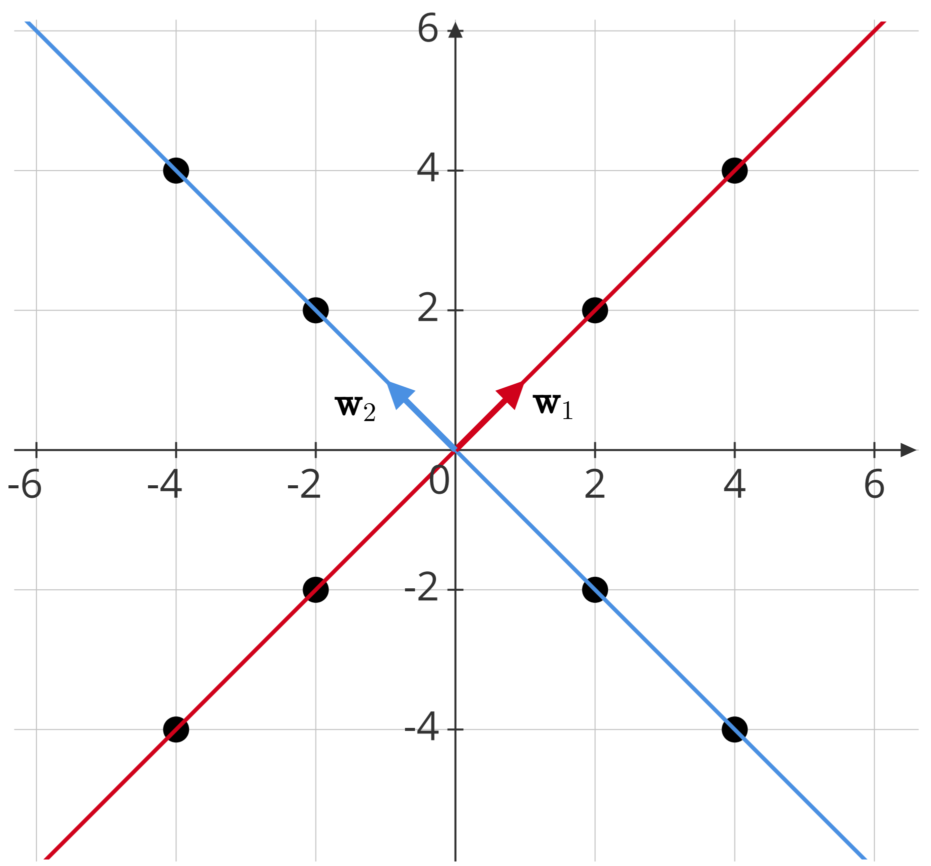
$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$



$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

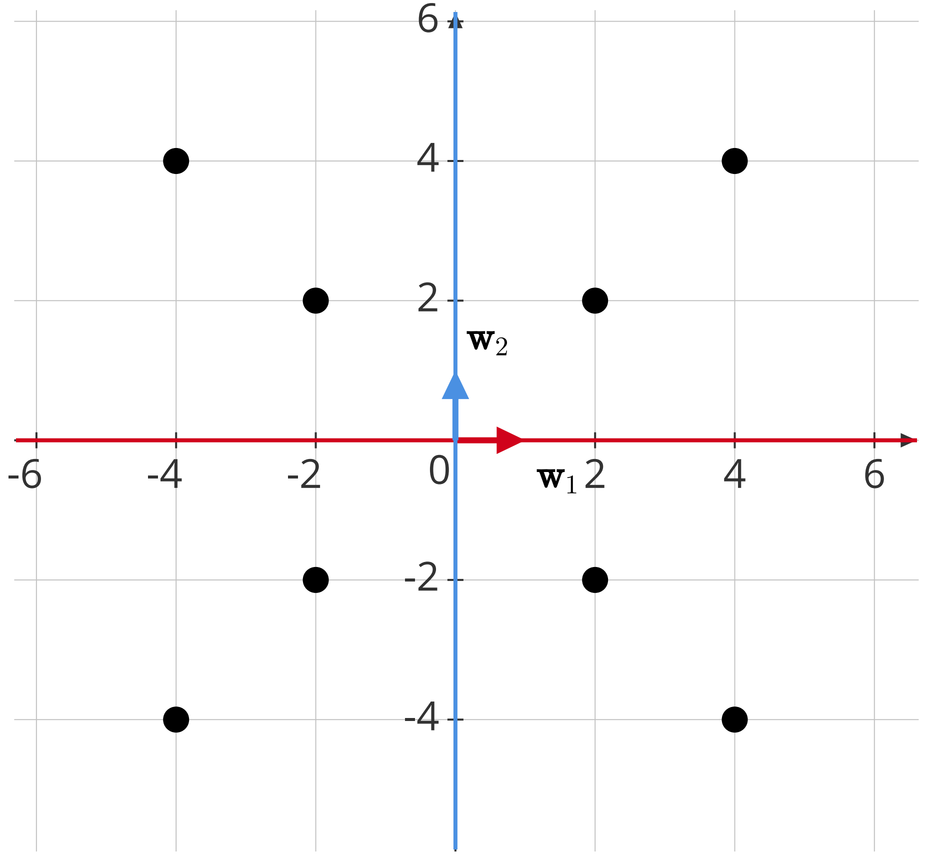
$$\mathbf{w}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Q-4

Find the principal components

$$\mathbf{X} = \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix}$$

$$\mathbf{X}^T = \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$



$$\mathbf{C} = \frac{1}{n} \cdot \mathbf{X}\mathbf{X}^T$$

$$= \frac{1}{8} \cdot \begin{bmatrix} 2 & -2 & 4 & -4 & 2 & -2 & 4 & -4 \\ 2 & -2 & 4 & -4 & -2 & 2 & -4 & 4 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -2 & -2 \\ 4 & 4 \\ -4 & -4 \\ 2 & -2 \\ -2 & 2 \\ 4 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \qquad \mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20$, $\sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20$, $\sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20$, $\sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$\nabla L(\mathbf{w}, D)$$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20,\ \sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$\nabla L(\mathbf{w}, D) = 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i$$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20,\ \sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$\begin{aligned}\nabla L(\mathbf{w}, D) &= 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i \\ &= 2 \cdot \sum_{i=1}^{n} [(\mathbf{w}^T \mathbf{u})c_i^2 - c_i y_i]\mathbf{u}\end{aligned}$$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20$, $\sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$\begin{aligned}\nabla L(\mathbf{w}, D) &= 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i \\ &= 2 \cdot \sum_{i=1}^{n} [(\mathbf{w}^T \mathbf{u})c_i^2 - c_i y_i]\mathbf{u} \\ &= 0\end{aligned}$$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20,\ \sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$
\begin{aligned}
\nabla L(\mathbf{w}, D) &= 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i \\
&= 2 \cdot \sum_{i=1}^{n} [(\mathbf{w}^T \mathbf{u})c_i^2 - c_i y_i]\mathbf{u} \\
&= 0
\end{aligned}
$$

$$\implies \mathbf{w}^T \mathbf{u}$$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20$, $\sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

$$\begin{aligned}
\nabla L(\mathbf{w}, D) &= 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i \\
&= 2 \cdot \sum_{i=1}^{n} [(\mathbf{w}^T \mathbf{u})c_i^2 - c_i y_i]\mathbf{u} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\implies \mathbf{w}^T \mathbf{u} &= \frac{\sum_{i=1}^{n} c_i y_i}{\sum_{i=1}^{n} c_i^2} \\
&= 0.2
\end{aligned}$$

Q-5

- $D = \{(c_1 \cdot \mathbf{u}, y_1),\ \cdots, (c_n \cdot \mathbf{u}, y_n)\}$
- $\mathbf{u} \in \mathbb{R}^d$ and $\mathbf{u} \neq 0$
- $c_i \neq 0$ for some i
- $\sum_{i=1}^{n} c_i \cdot y_i = 20,\ \sum_{i=1}^{n} c_i^2 = 100$
- $\mathbf{x}_{\texttt{test}} = 5 \cdot \mathbf{u}$

$$\begin{aligned} \nabla L(\mathbf{w}, D) &= 2 \cdot \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)\mathbf{x}_i \\ &= 2 \cdot \sum_{i=1}^{n} [(\mathbf{w}^T \mathbf{u})c_i^2 - c_i y_i]\mathbf{u} \\ &= 0 \end{aligned}$$

$$\begin{aligned} \implies \mathbf{w}^T \mathbf{u} &= \frac{\sum_{i=1}^{n} c_i y_i}{\sum_{i=1}^{n} c_i^2} \\ &= 0.2 \end{aligned}$$

$$\begin{aligned} \mathbf{w}^T \mathbf{x}_{\texttt{test}} &= \mathbf{w}^T (5 \cdot \mathbf{u}) \\ &= 5 \cdot \left(\mathbf{w}^T \mathbf{u}\right) \\ &= 1 \end{aligned}$$

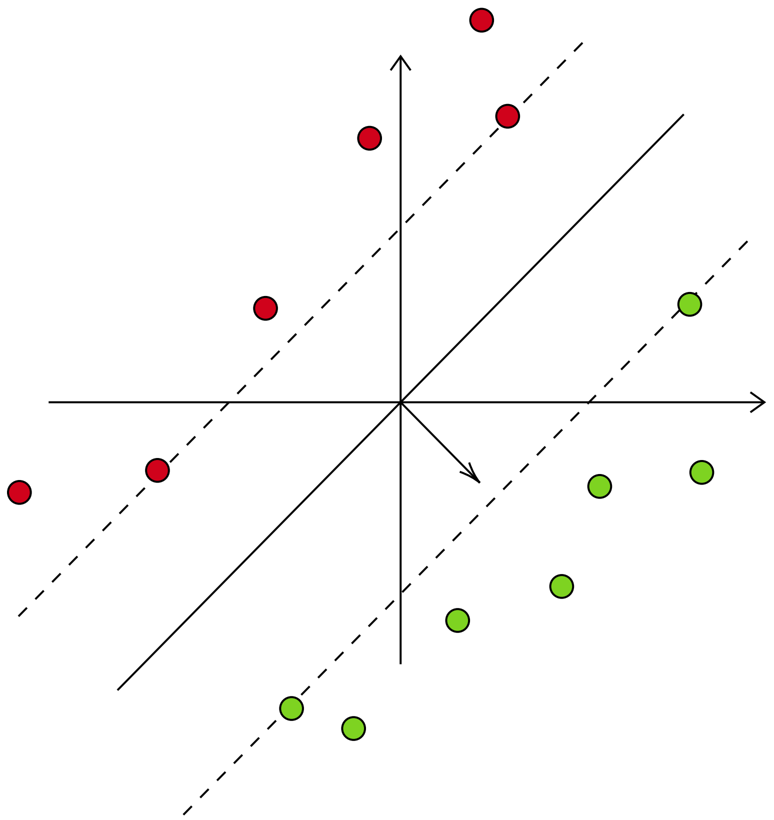
$$L(\mathbf{w}, D) = \sum_{i=1}^{n} (\mathbf{w}^T \mathbf{x}_i - y_i)^2$$

Q-6

- Hard-margin, linear-SVM
- Add 10 new + data-points
  - correct side of margin
  - very far away from it
- Retrain the classifier
- Comment on decision boundary

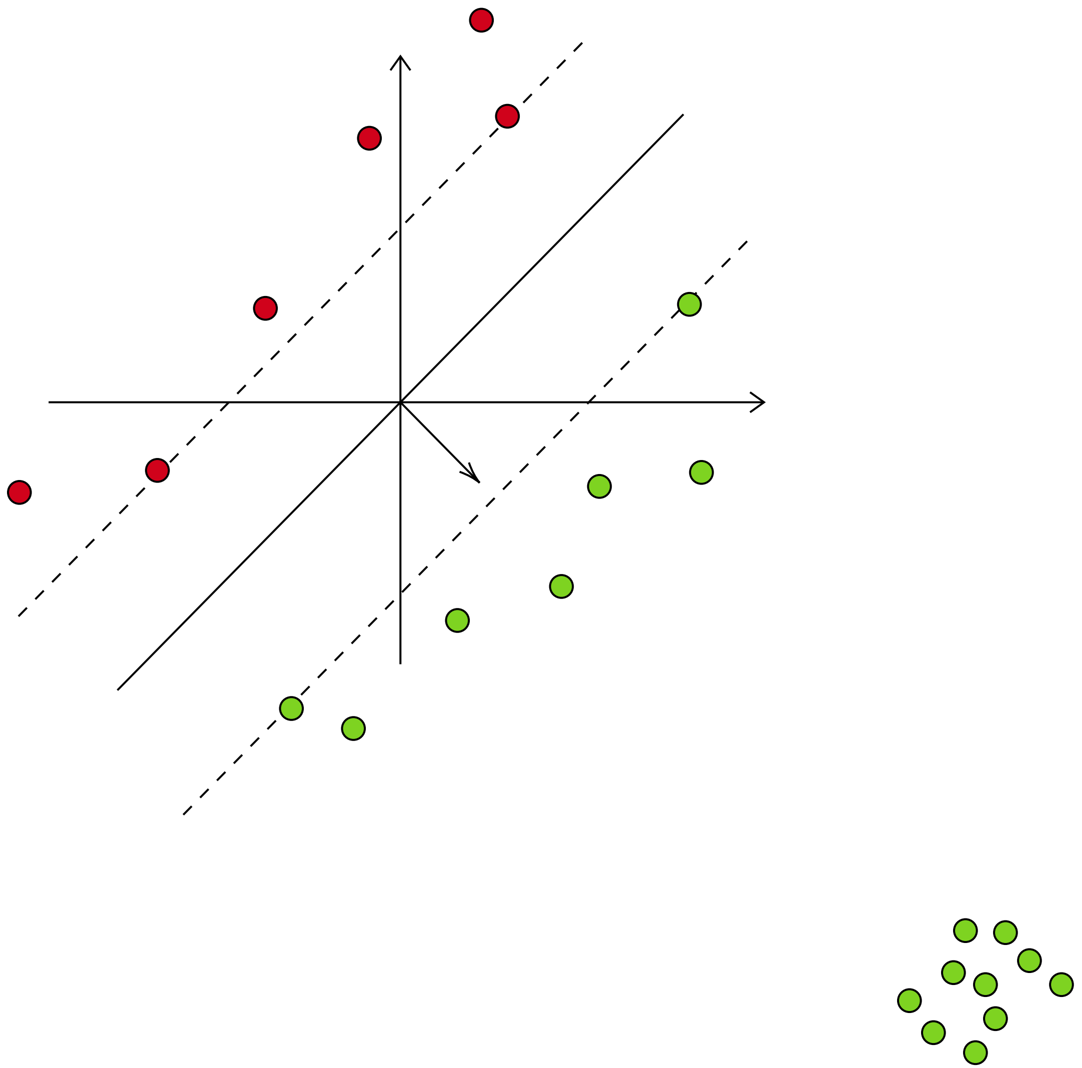
Q-6

- Hard-margin, linear-SVM
- Add 10 new $+$ data-points
  - correct side of margin
  - very far away from it
- Retrain the classifier
- Comment on decision boundary



Q-6

- Hard-margin, linear-SVM
- Add 10 new + data-points
  - correct side of margin
  - very far away from it
- Retrain the classifier
- Comment on decision boundary

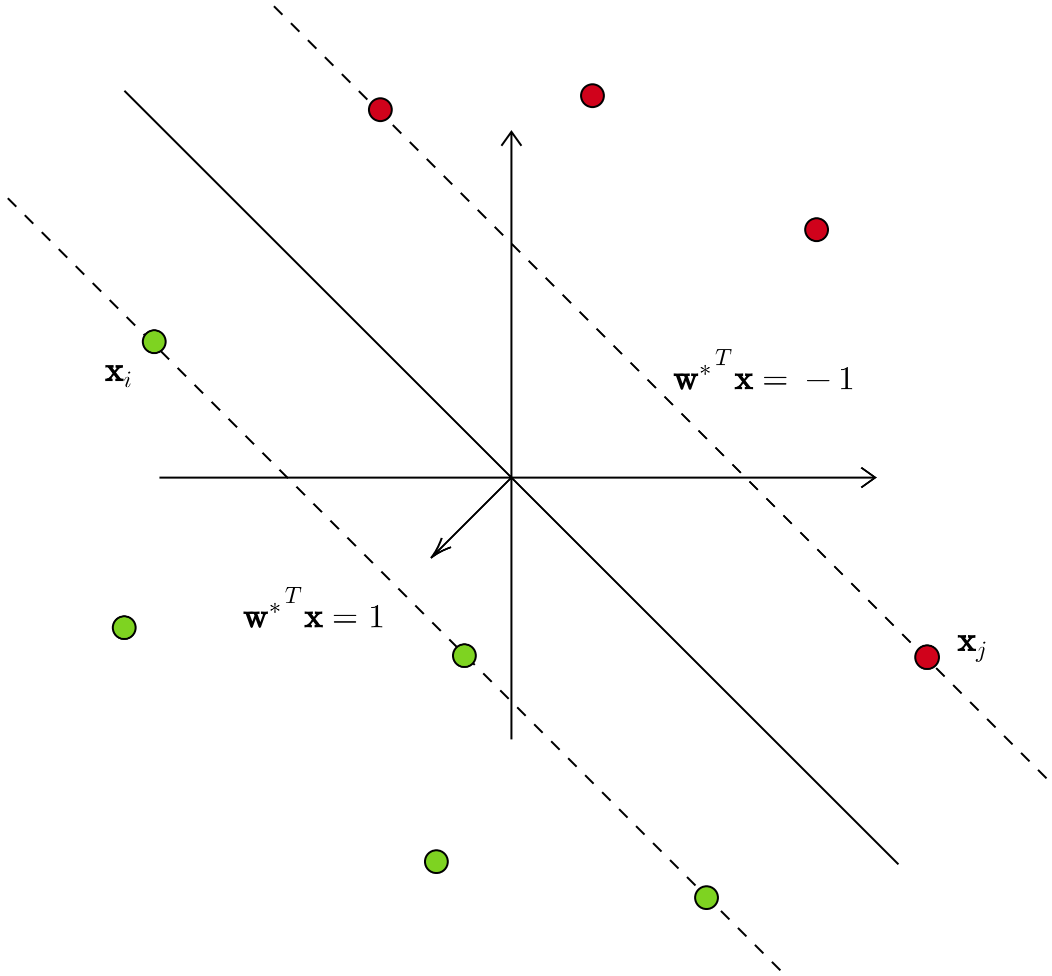


Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary

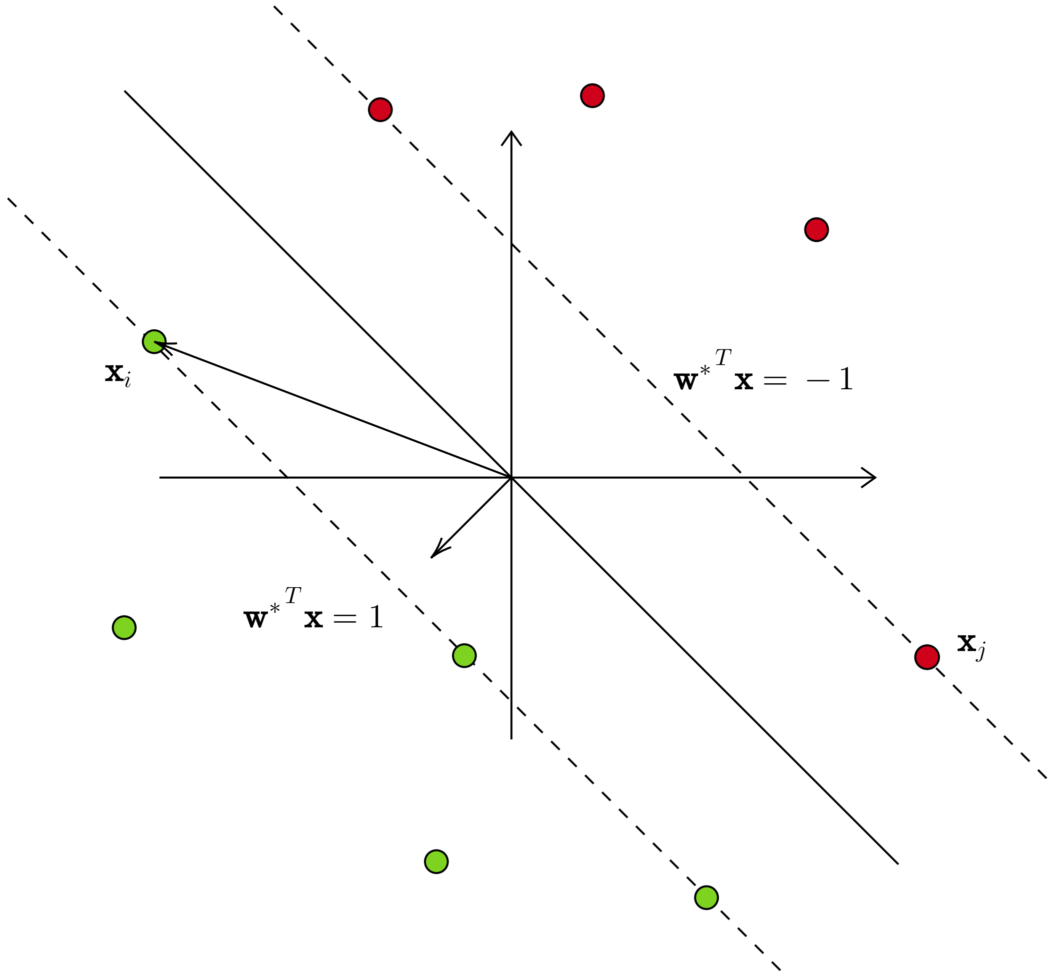
Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



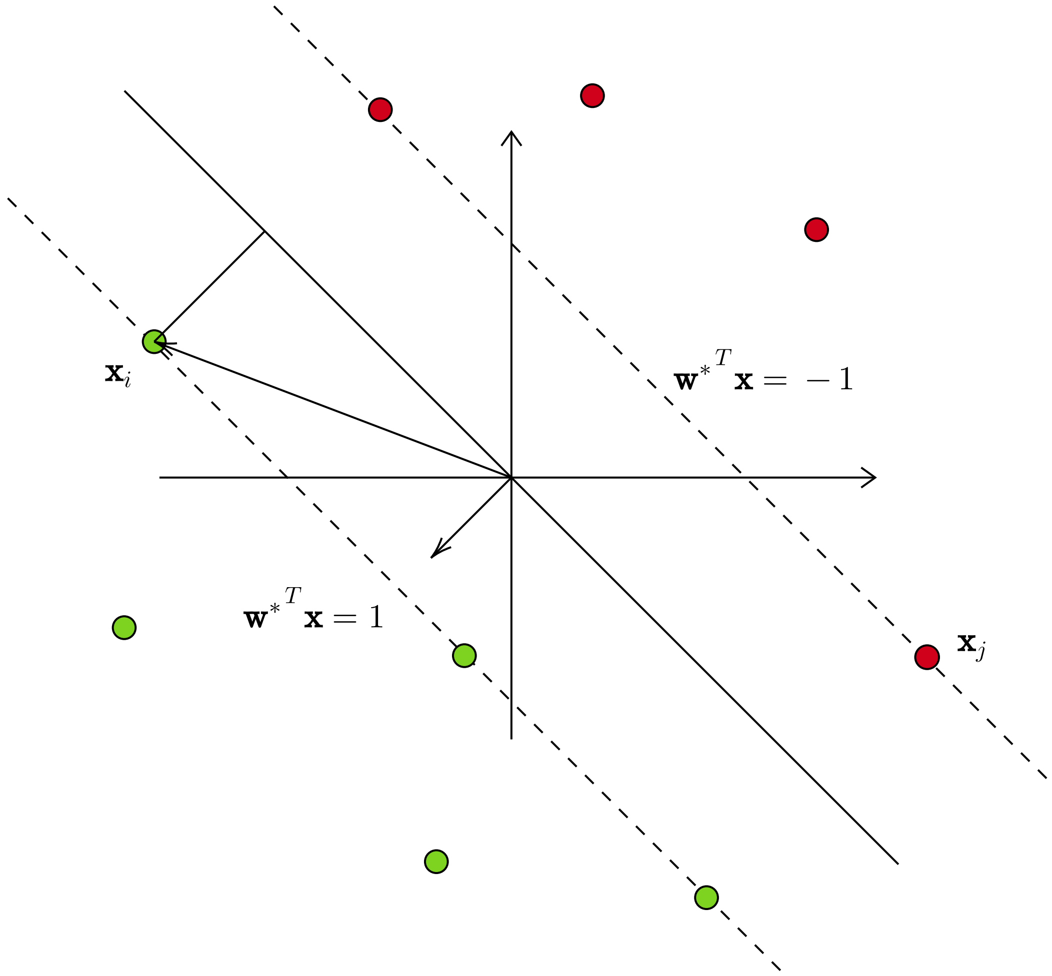
Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



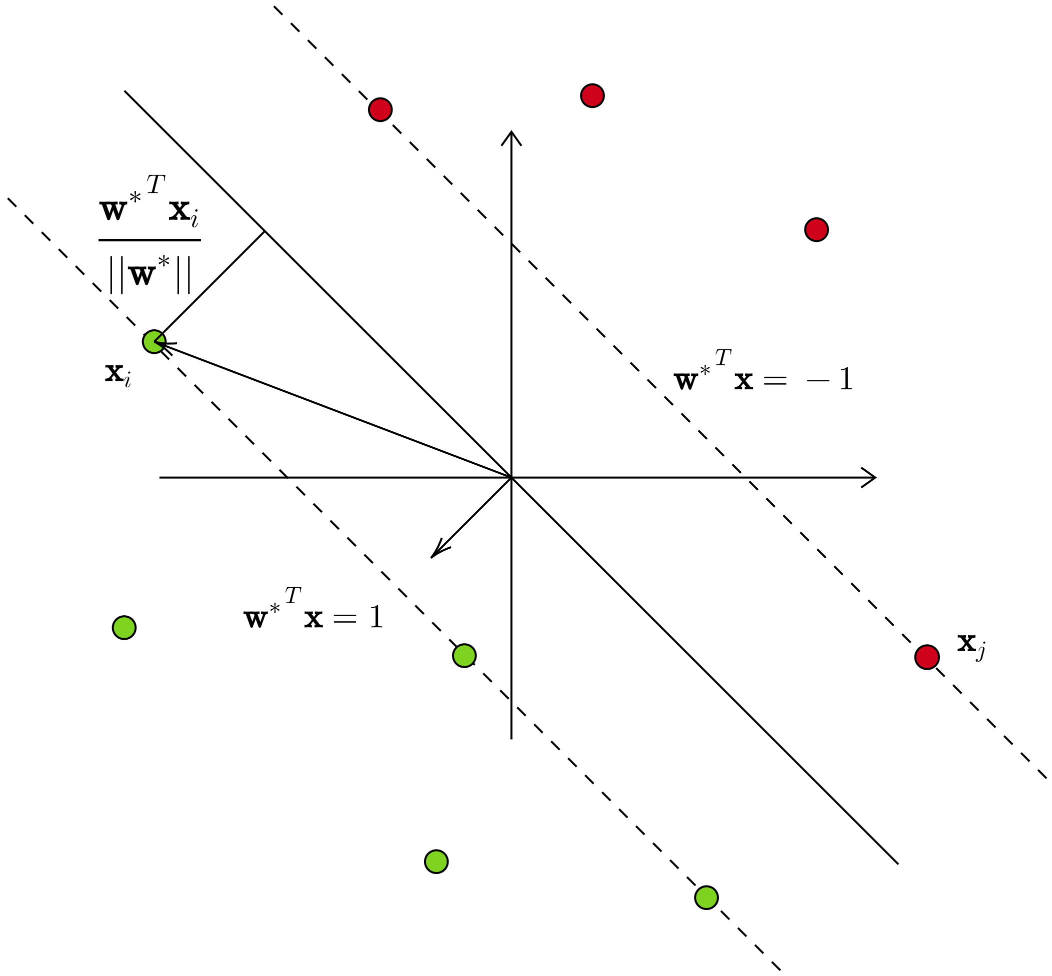
Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



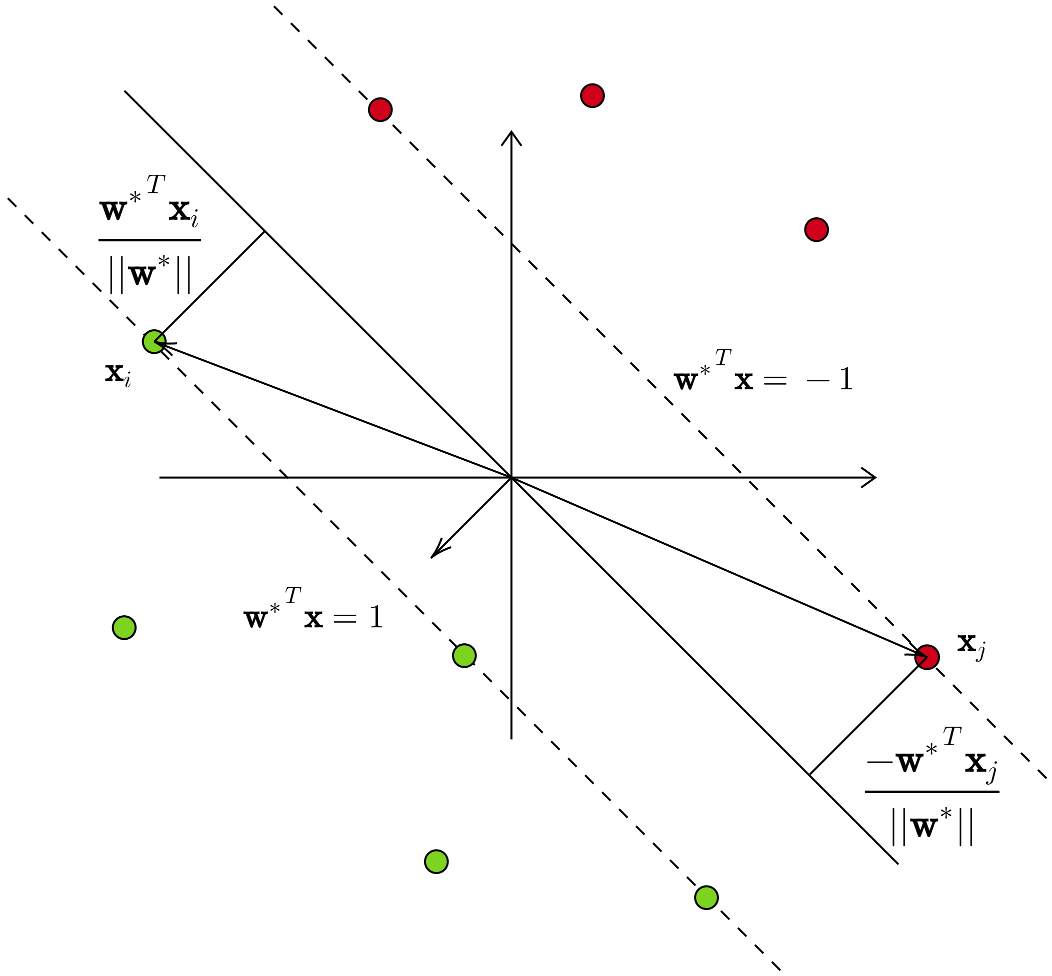
Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



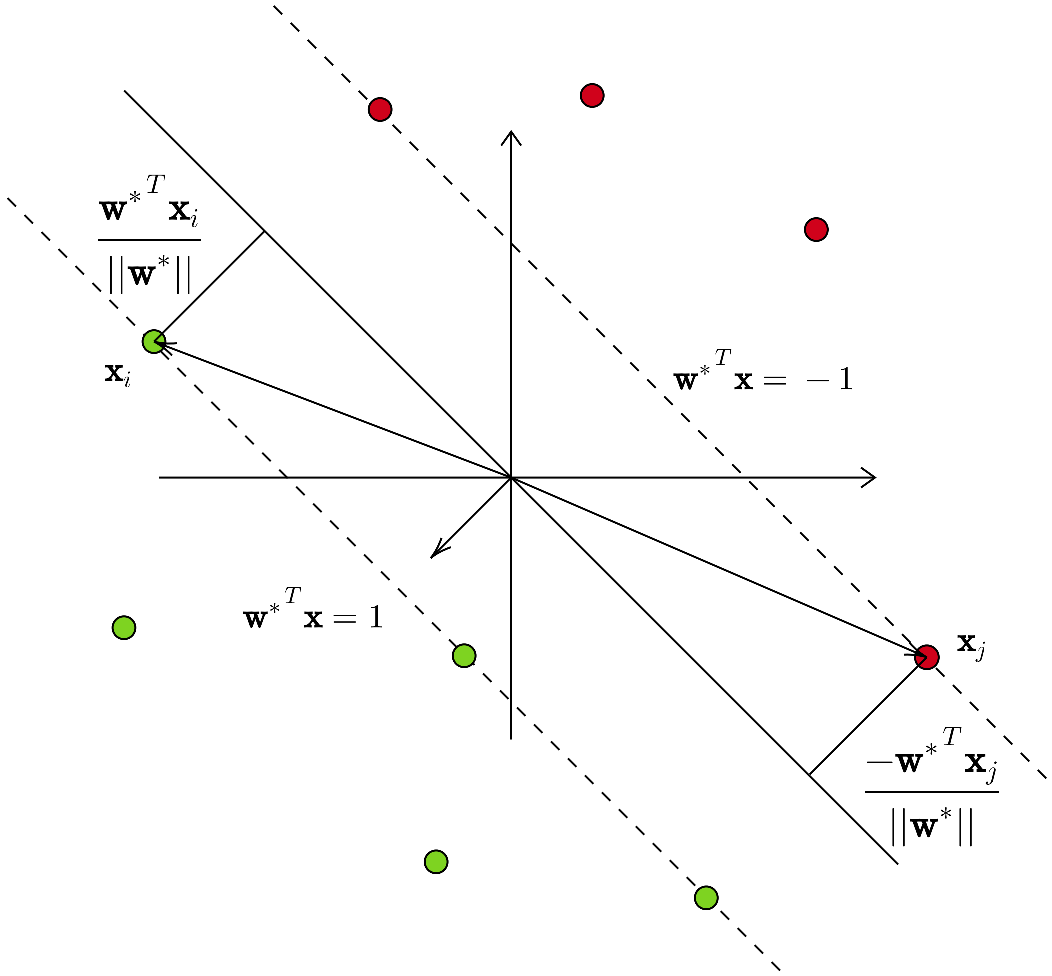
Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



Q-7

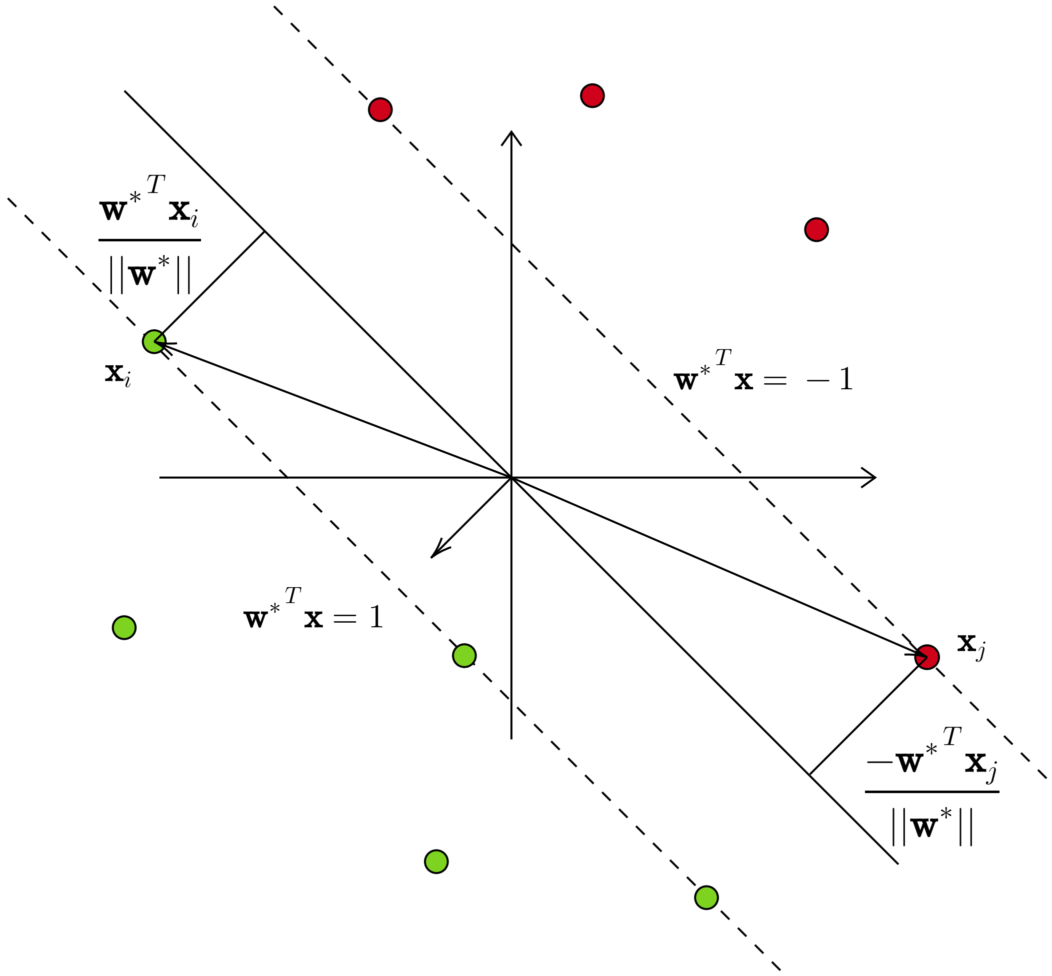
- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



$$d =$$

Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary



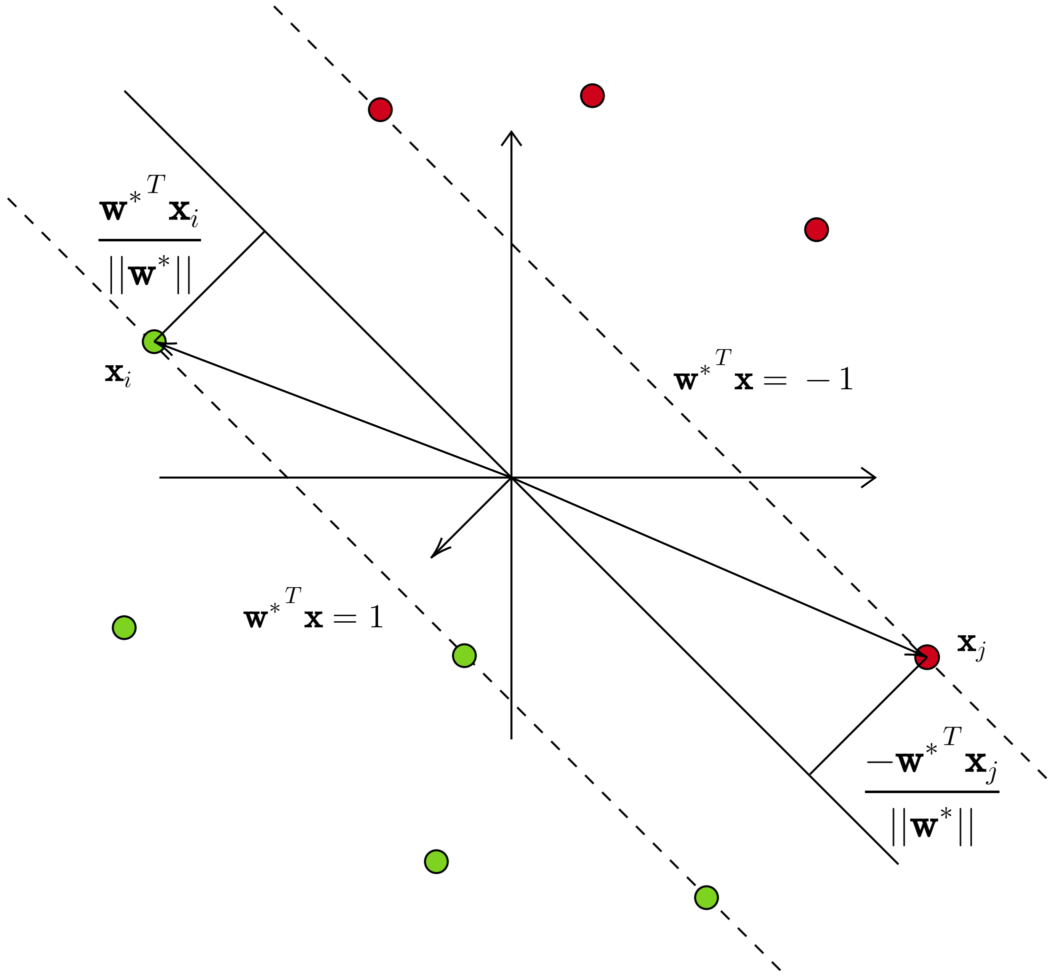
$$d = \frac{1}{||\mathbf{w}^*||} = \frac{1}{\sqrt{1^2 + 2^2 + 3^2}} = \frac{1}{\sqrt{14}}$$

Q-7

- Hard-margin, linear-SVM
- Optimal weight vector $\mathbf{w}^* = [1 \quad 2 \quad 3]^T$
- Distance of the closest point from the boundary

Geometric margin

$$d = \frac{1}{||\mathbf{w}^*||} = \frac{1}{\sqrt{1^2 + 2^2 + 3^2}} = \frac{1}{\sqrt{14}}$$



Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \mathtt{diag}(r_1,\ \cdots .\, r_n)$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1, \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1,\ \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\nabla L(\mathbf{w})$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1, \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\nabla L(\mathbf{w}) = \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1, \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\begin{aligned}\nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i\end{aligned}$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1,\ \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{aligned} \nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i \end{aligned}$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1,\ \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{aligned} \nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i \end{aligned}$$

$$\mathbf{XR} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{bmatrix}$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \mathtt{diag}(r_1, \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix}, \ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{aligned} \nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i \end{aligned}$$

$$\begin{aligned} \mathbf{XR} &= \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{bmatrix} \\ &= \begin{bmatrix} | & & | \\ r_1 \mathbf{x}_1 & \cdots & r_n \mathbf{x}_n \\ | & & | \end{bmatrix} \end{aligned}$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \texttt{diag}(r_1,\ \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{aligned} \nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i \end{aligned}$$

$$\begin{aligned} \mathbf{XR} &= \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{bmatrix} \\ &= \begin{bmatrix} | & & | \\ r_1 \mathbf{x}_1 & \cdots & r_n \mathbf{x}_n \\ | & & | \end{bmatrix} \end{aligned}$$

$$\mathbf{XRX}^T \mathbf{w} = \mathbf{XRy}$$

Q-8

- Weighted linear regression
- Weight for each data-point in the loss function
- $\mathbf{R} = \mathtt{diag}(r_1,\ \cdots . r_n)$

$$L(\mathbf{w}) = \frac{1}{2} \cdot \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right)^2$$

$$\mathbf{X} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{aligned} \nabla L(\mathbf{w}) &= \sum_{i=1}^{n} r_i \left(\mathbf{w}^T \mathbf{x}_i - y_i\right) \mathbf{x}_i \\ &= \sum_{i=1}^{n} (r_i \mathbf{x}_i) \mathbf{x}_i^T \mathbf{w} - \sum_{i=1}^{n} r_i y_i \mathbf{x}_i \end{aligned}$$

$$\mathbf{XR} = \begin{bmatrix} | & & | \\ \mathbf{x}_1 & \cdots & \mathbf{x}_n \\ | & & | \end{bmatrix} \begin{bmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ r_1 \mathbf{x}_1 & \cdots & r_n \mathbf{x}_n \\ | & & | \end{bmatrix}$$

$$\mathbf{XRX}^T \mathbf{w} = \mathbf{XRy}$$

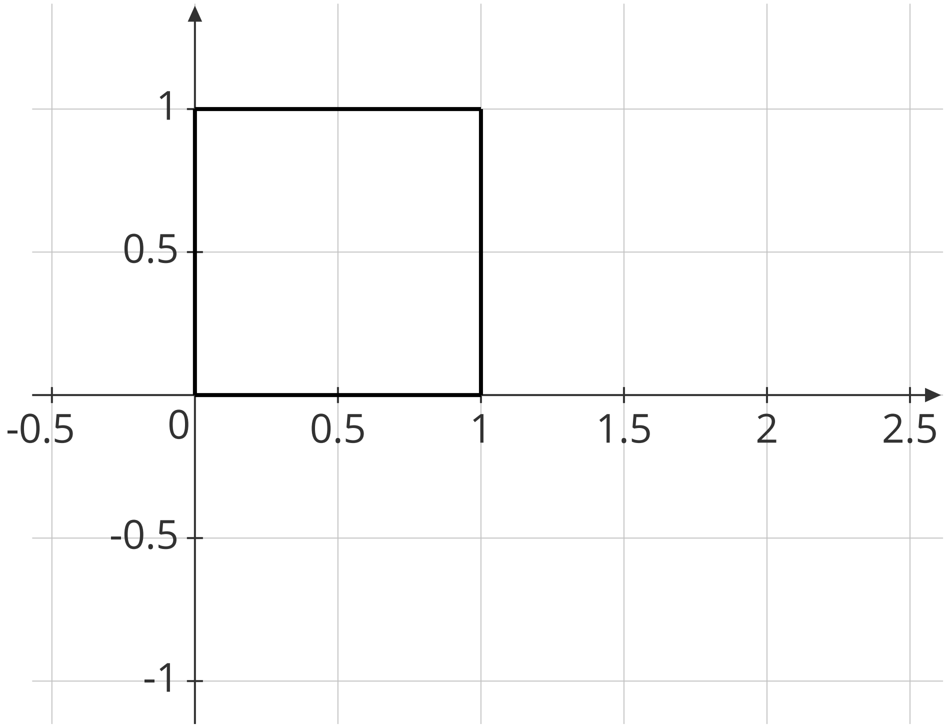
$$\mathbf{w} = \left(\mathbf{XRX}^T\right)^{-1} \mathbf{XRy}$$

Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1

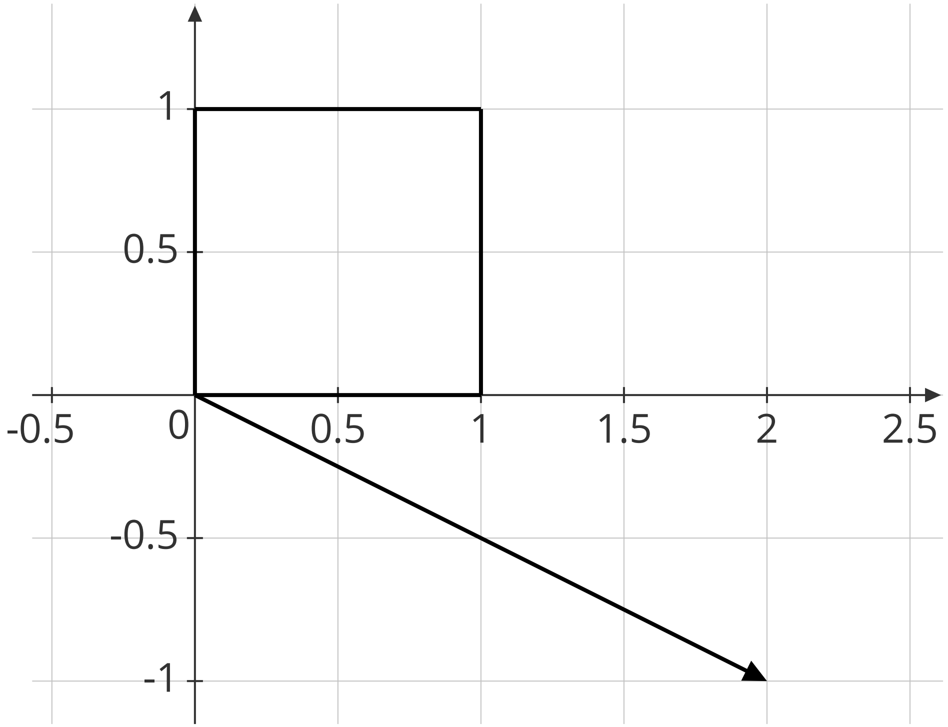
Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1



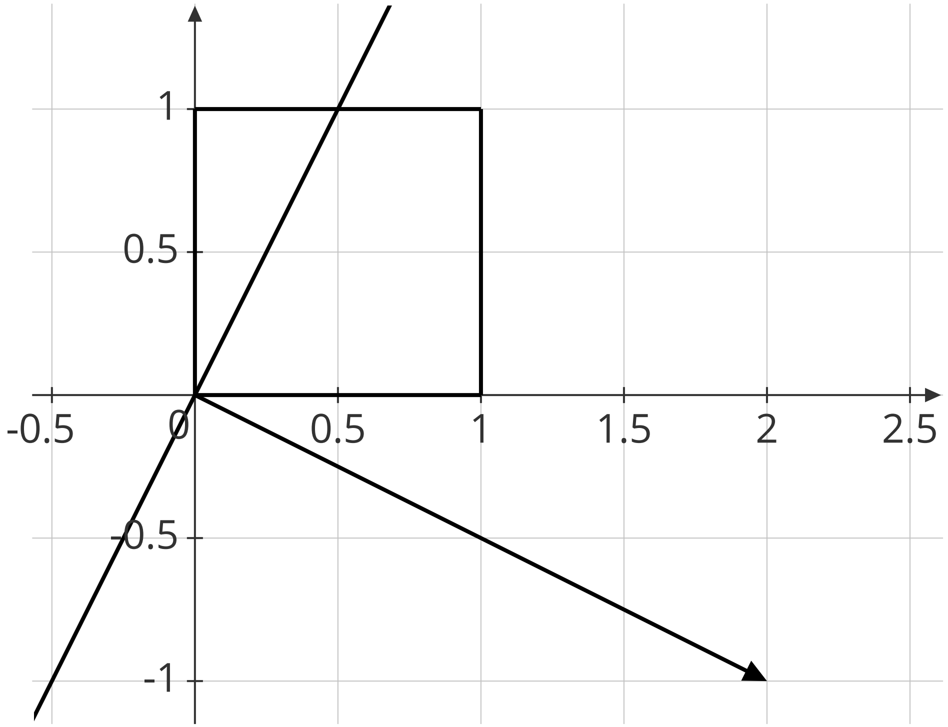
Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1



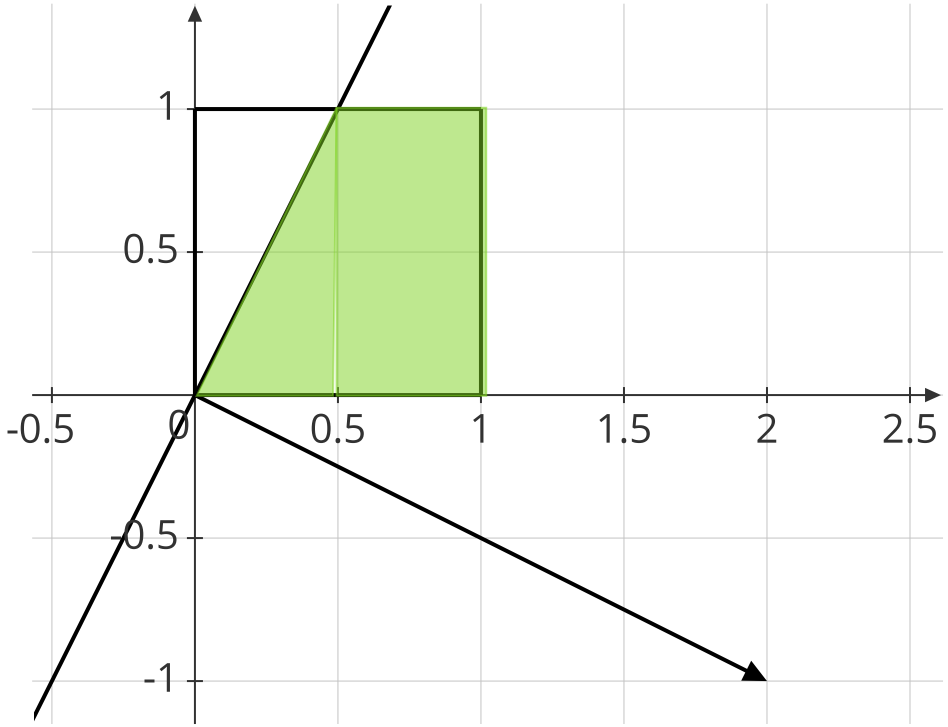
Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1



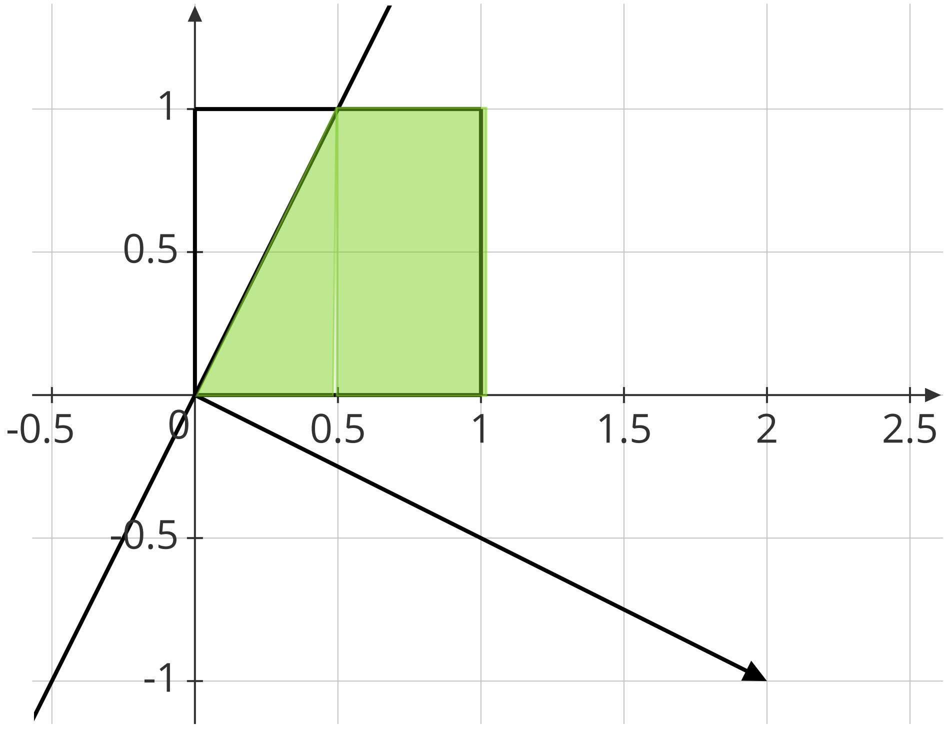
Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1



Q-9

- Hard-margin, linear-SVM
- $\mathbf{w}^* = [2 \quad -1]^T$
- Unit square (0,0), (1, 0), (0, 1), (1, 1)
- Probability that a point picked at random from the unit square is is predicted as 1



$$P(y = 1 \mid \mathbf{x}) = 1 - \frac{1}{2} \cdot 1 \cdot \frac{1}{2}$$

$$= 1 - 0.25$$

$$= 0.75$$

Q-10

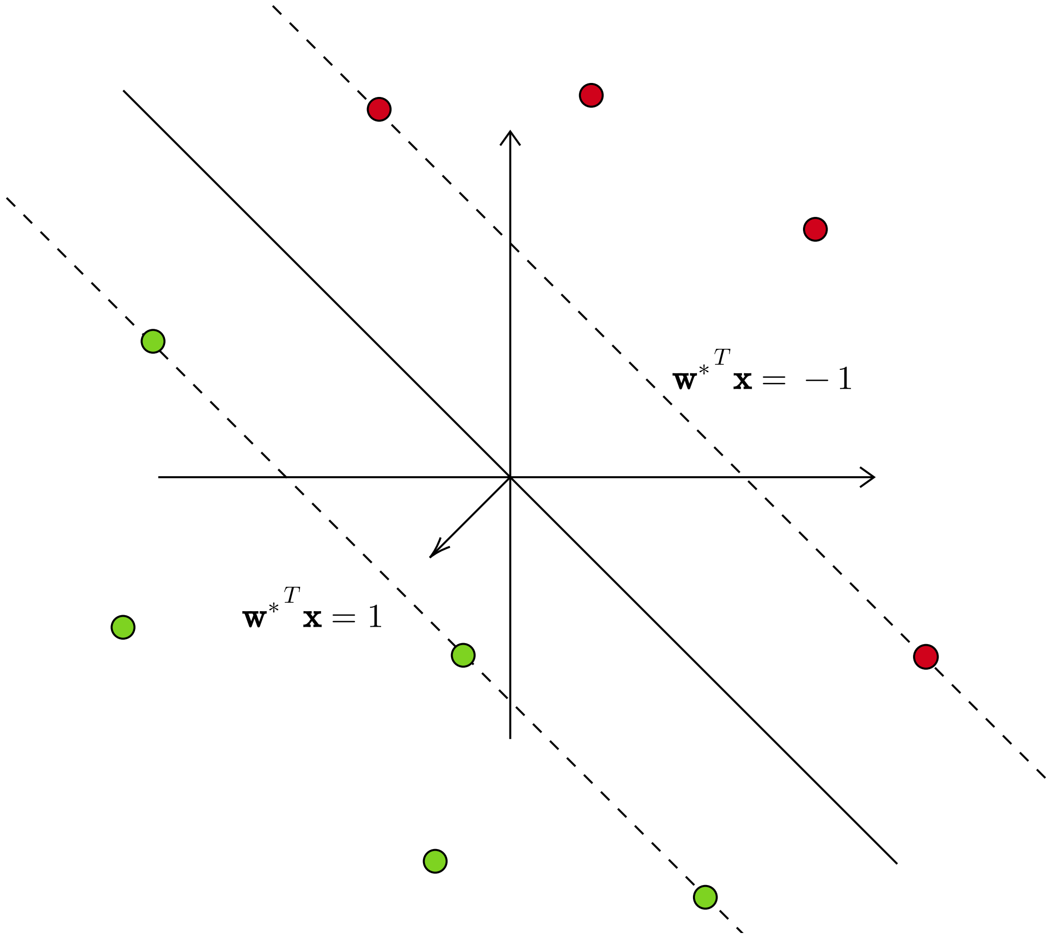
Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.

Q-10

Select all true statements regarding a hard-margin SVM:

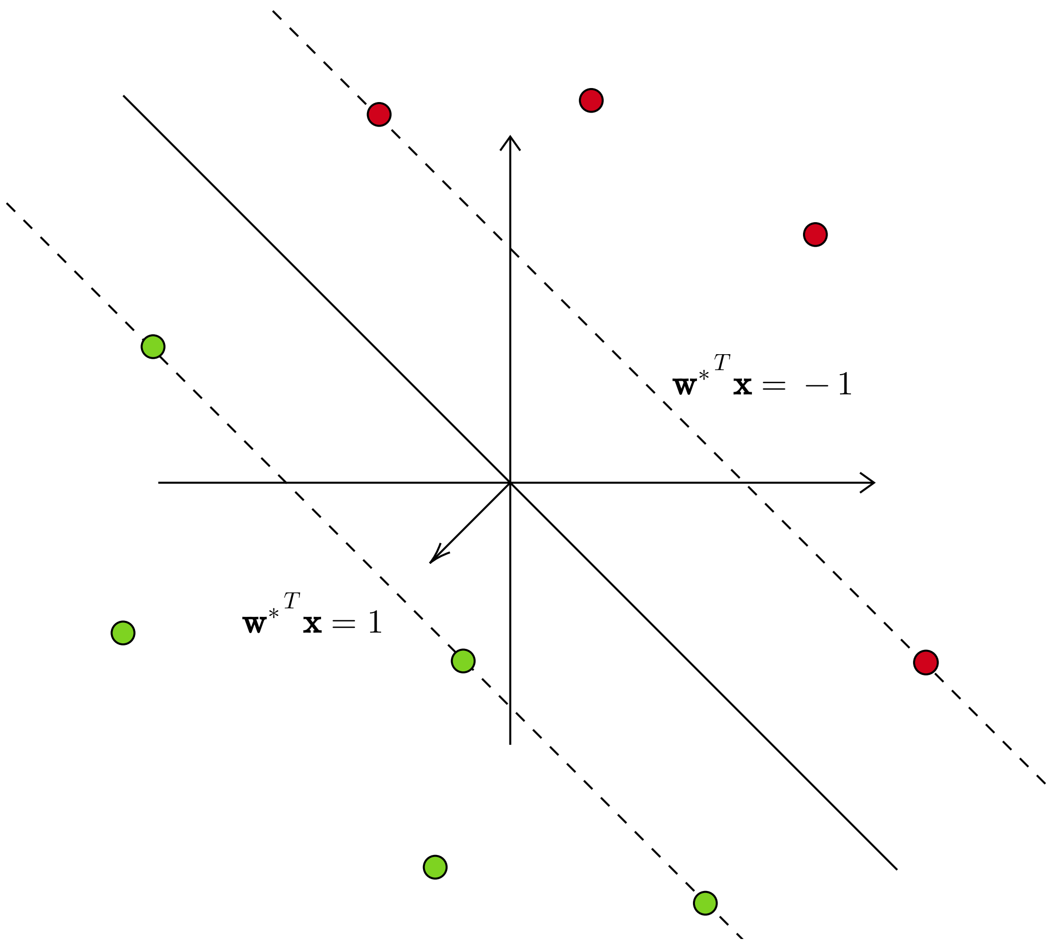
(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.

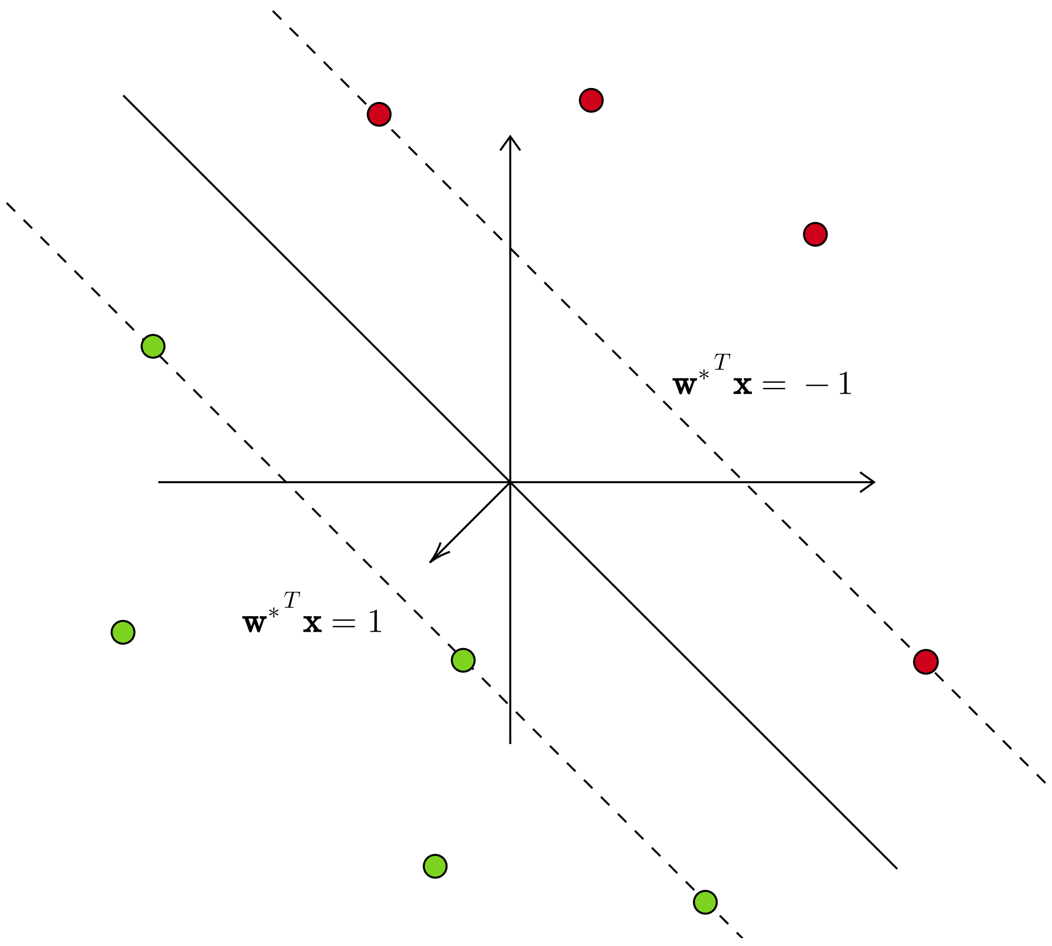


$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \geqslant 1 \Longrightarrow 1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \leqslant 0$$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



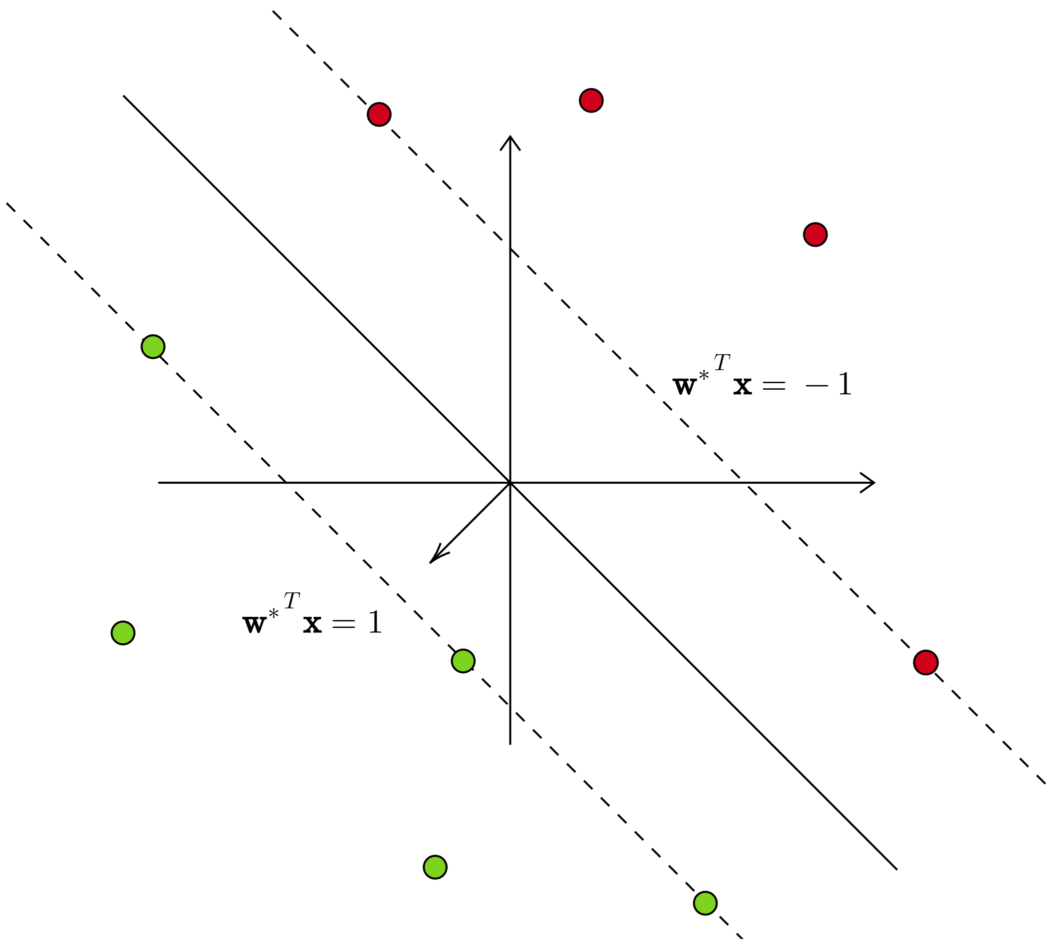
$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \geqslant 1 \Longrightarrow 1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \leqslant 0$$

$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i\right] = 0$$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



$$\left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i \geqslant 1 \implies 1 - \left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i \leqslant 0$$

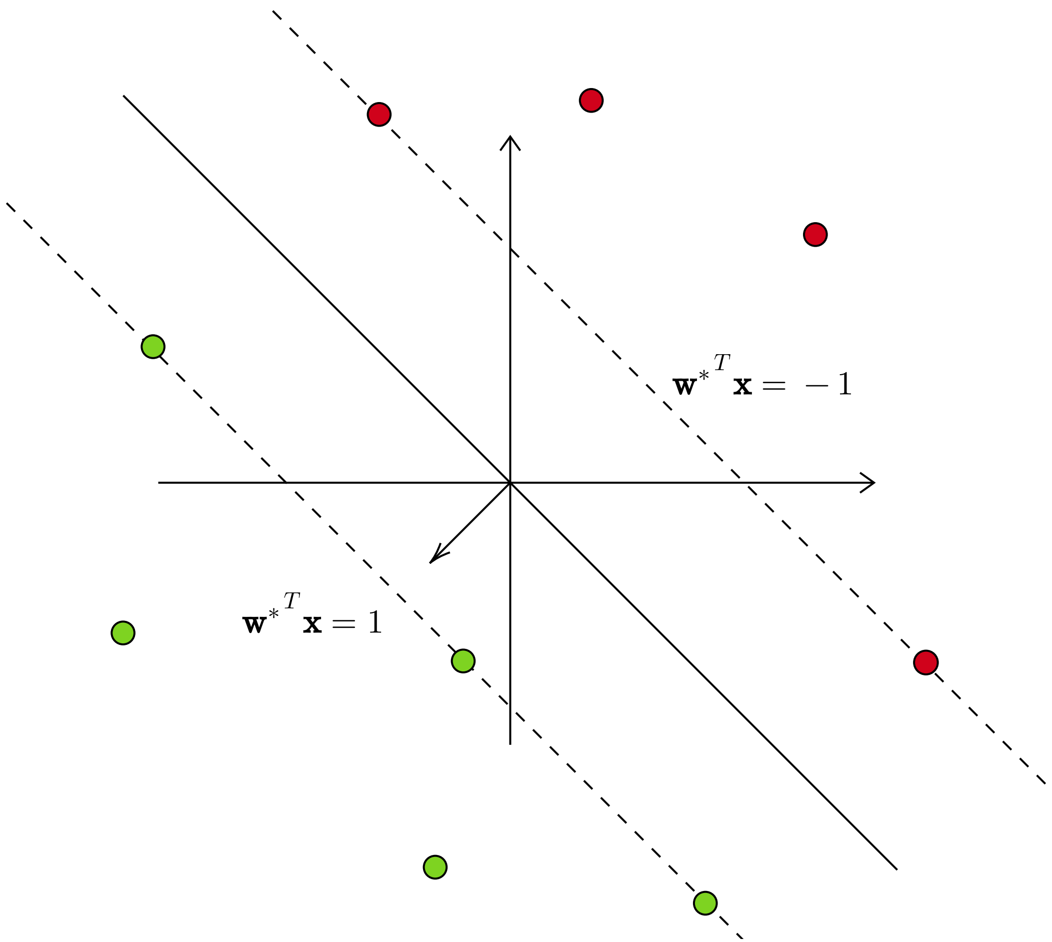
$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i\right] = 0$$

Support vector $\implies$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \geqslant 1 \implies 1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \leqslant 0$$

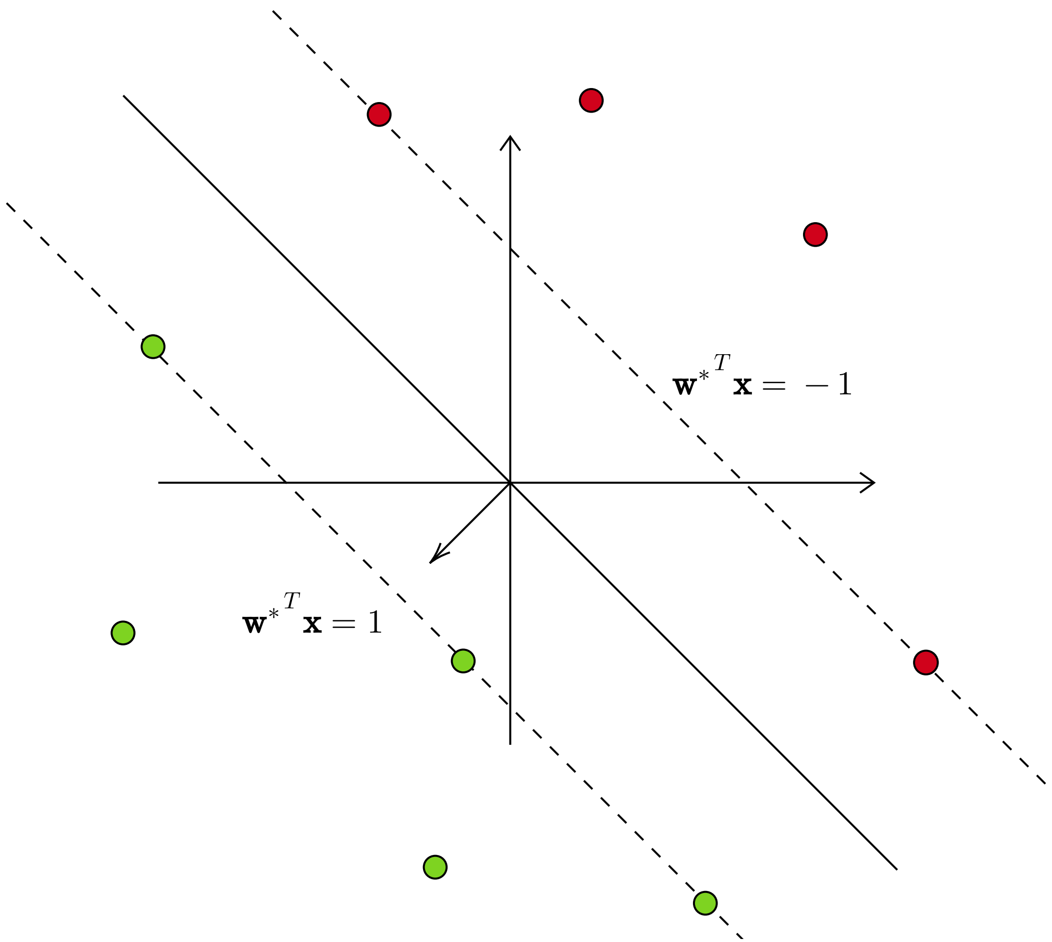
$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i\right] = 0$$

Support vector $\implies$ $\alpha_i^* > 0$ $\implies$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



$$\left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i \geqslant 1 \implies 1 - \left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i \leqslant 0$$

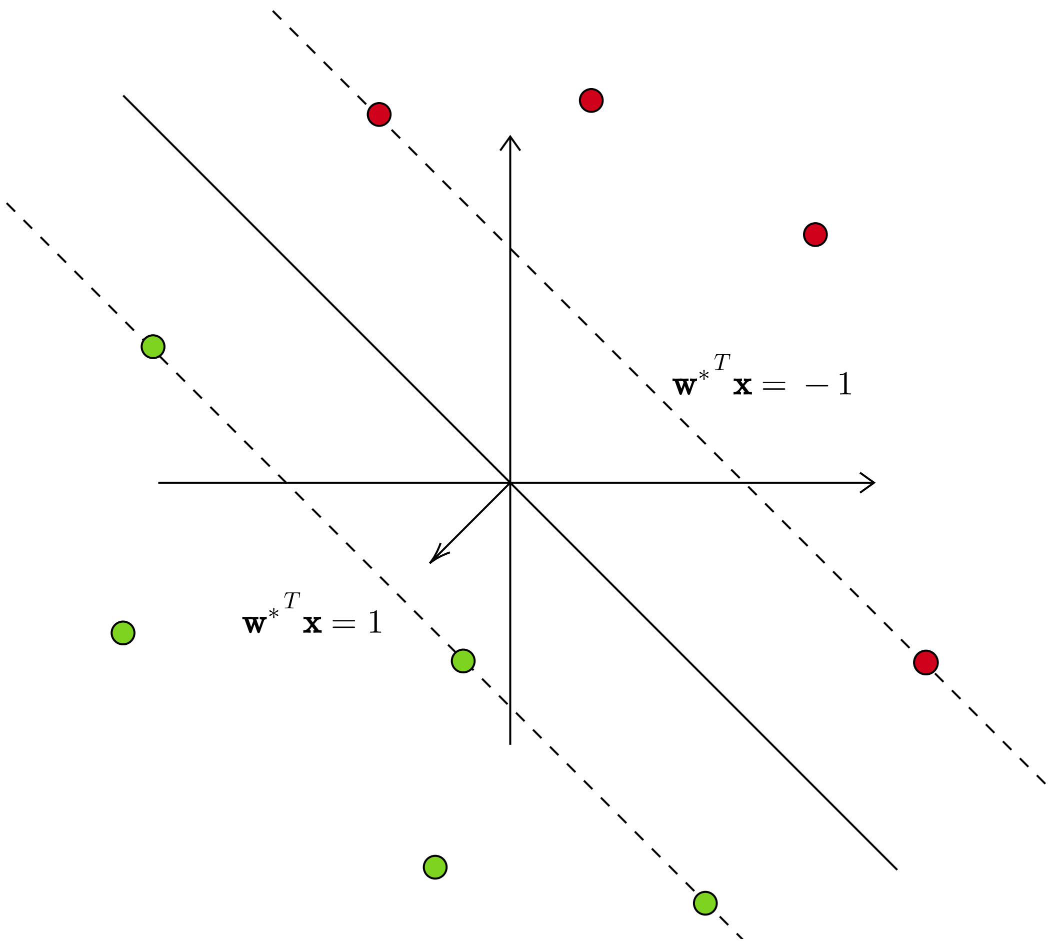
$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i\right] = 0$$

Support vector $\implies \alpha_i^* > 0 \implies \left(\mathbf{w^*}^T\mathbf{x}_i\right)y_i = 1$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \geqslant 1 \Longrightarrow 1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \leqslant 0$$

$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i\right] = 0$$

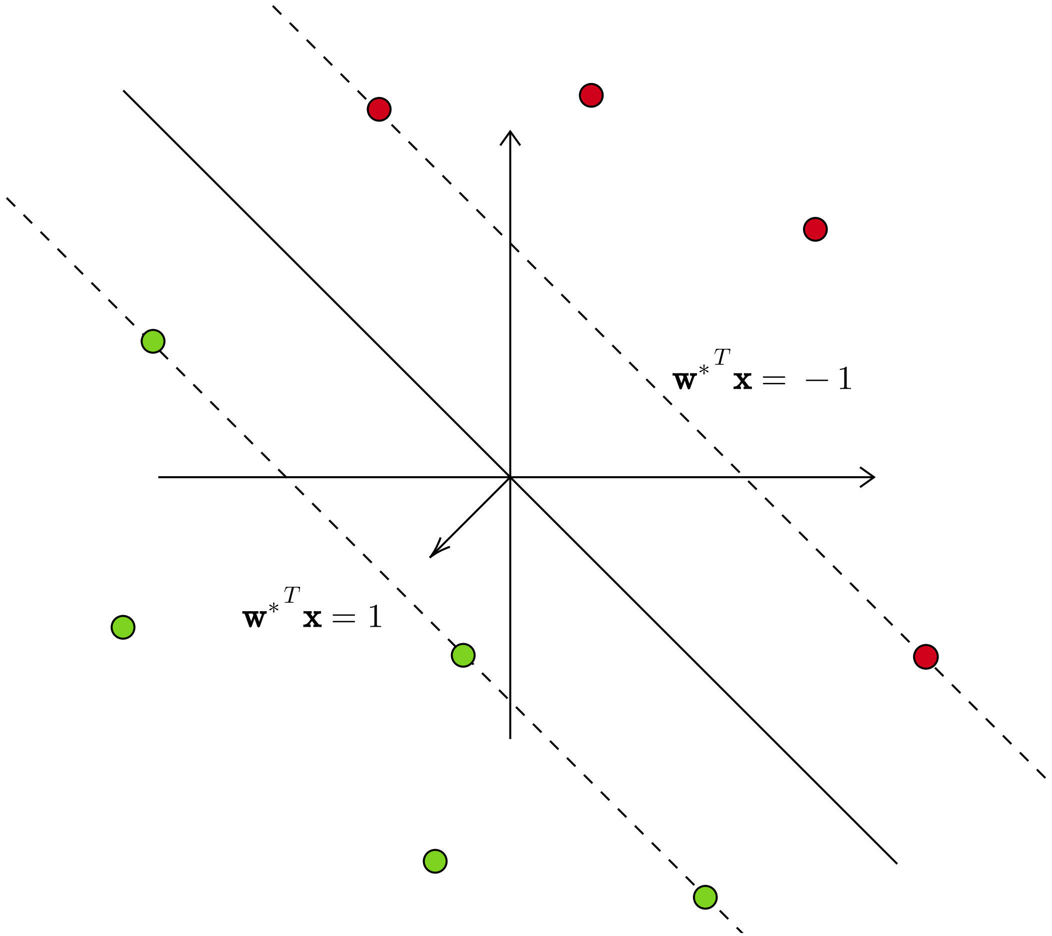
Support vector $\Longrightarrow \alpha_i^* > 0 \Longrightarrow \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i = 1$

$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i = 1$$

Q-10

Select all true statements regarding a hard-margin SVM:

(1) Every support vector lies on one of the two supporting hyperplanes.
(2) Every point on one of the two supporting hyperplanes is a support vector.



$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \geqslant 1 \Longrightarrow 1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i \leqslant 0$$

$$\alpha_i^* \cdot \left[1 - \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i\right] = 0$$

Support vector $\Longrightarrow \alpha_i^* > 0 \Longrightarrow \left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i = 1$

$$\left(\mathbf{w}^{*T}\mathbf{x}_i\right)y_i = 1 \quad \not\Longrightarrow \quad \alpha_i^* > 0$$