



# Παραγωγική Τεχνητή Νοημοσύνη: Generative AI

---

Κωνσταντίνος Καραμανής

The University of Texas at Austin & Archimedes/Athena RC

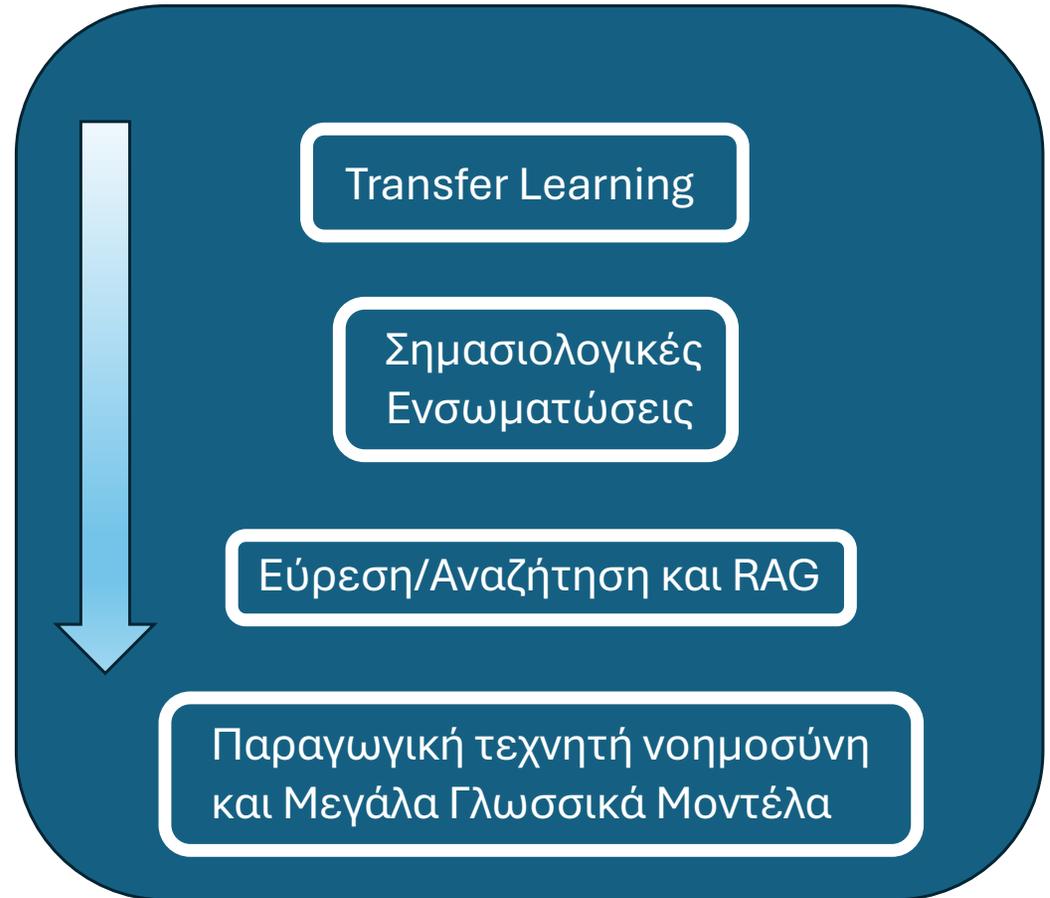
[constantine@utexas.edu](mailto:constantine@utexas.edu)

<https://caramanis.github.io/>





# Η Γέφυρα προς την Παραγωγική Τεχνητή Νοημοσύνη



# Λογιστική Παλινδρόμηση

---

- Γεωμετρική εξήγηση της λογιστικής παλινδρόμησης
- Εσωτερικό Γινόμενο
- Οι βασικές ιδέες των **embeddings** (ενσωματώσεων)



Ας θυμηθούμε τα  
προηγούμενα...



# Λογιστική Παλινδρόμηση

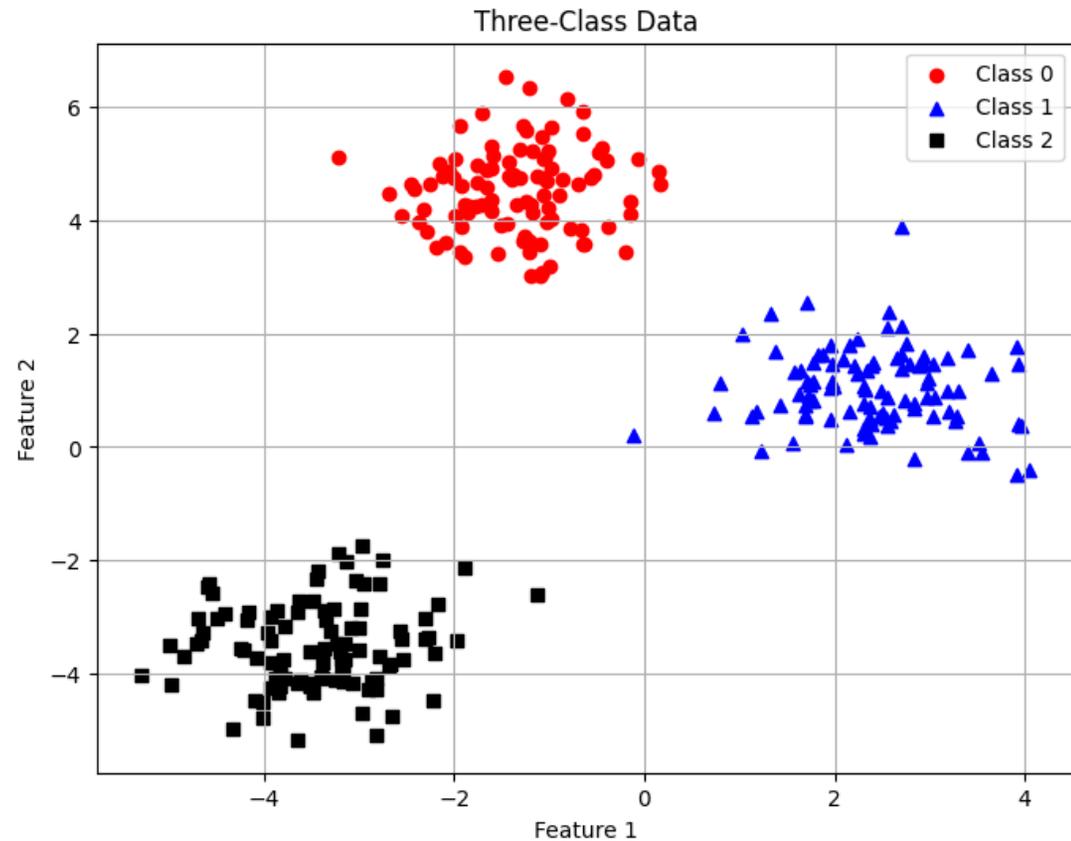
---

- Γραμμικός διαχωρισμός του χώρου
- Δεν ελαχιστοποιεί τα «λάθη»
- Ελαχιστοποιεί την «απώλεια» του γραμμικού διαχωρισμού
- **Είναι το πιο απλό νευρωνικό δίκτυο**

# Λογιστική Παλινδρόμηση: Παραδείγματα

---

- Τρεις κλάσεις/κατηγορίες
- Δύο δεδομένα εισόδου (features)

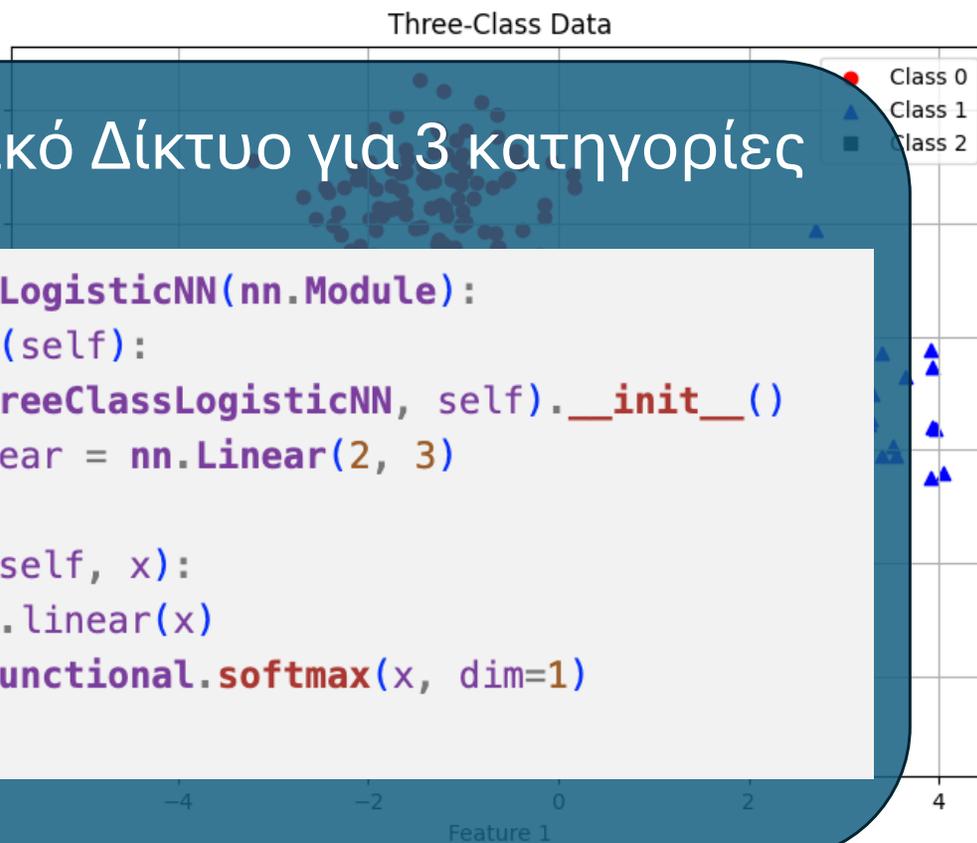


# Λογιστική Παλινδρόμηση: Παραδείγματα

## Νευρωνικό Δίκτυο για 3 κατηγορίες

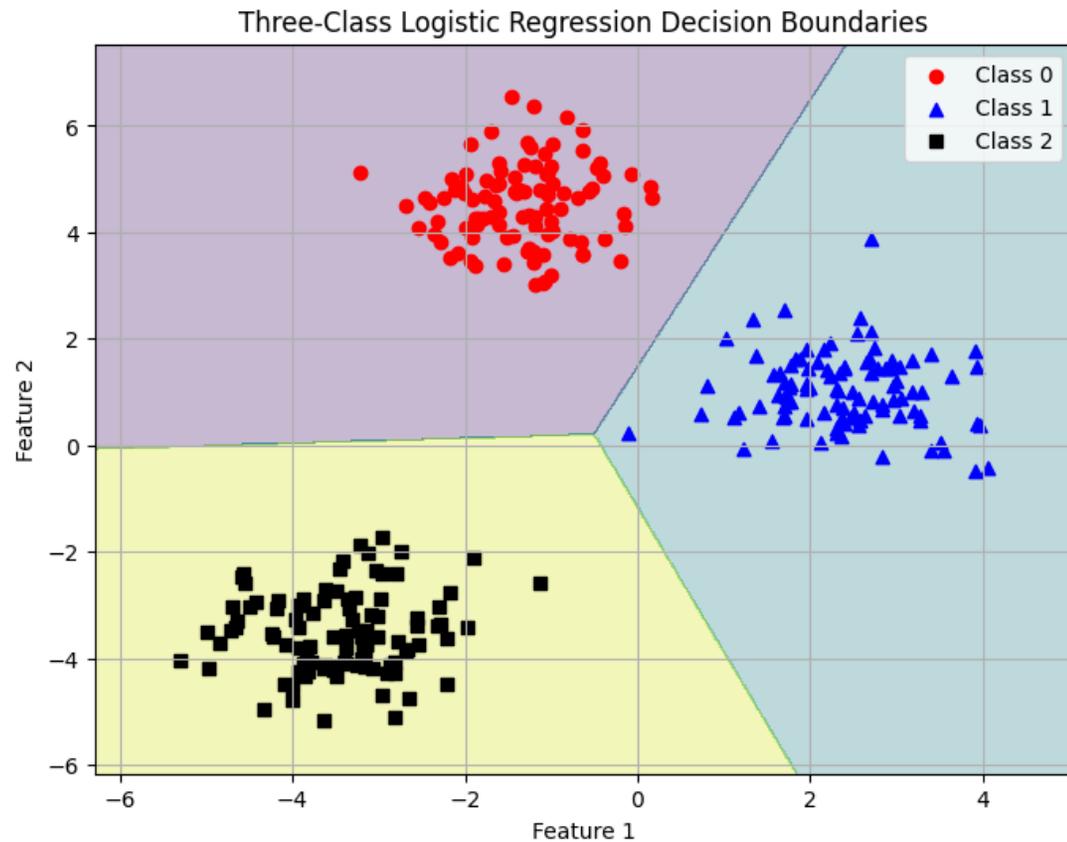
- Τρεις κλάσεις/κατηγορίες
- Δύο δεδομένα εισόδου (feature 1, feature 2)

```
class ThreeClassLogisticNN(nn.Module):  
    def __init__(self):  
        super(ThreeClassLogisticNN, self).__init__()  
        self.linear = nn.Linear(2, 3)  
  
    def forward(self, x):  
        x = self.linear(x)  
        x = nn.functional.softmax(x, dim=1)  
        return x
```



# Λογιστική Παλινδρόμηση: Παραδείγματα

- Τρεις κλάσεις/κατηγορίες
- Δύο δεδομένα εισόδου (features)



Κωνσταντίνος Καραμανής

# Λογιστική Παλινδρόμηση: Τι κάνει;

- Ο κώδικάς μας εξηγεί ακριβώς τι κάνει η λογιστική παλινδρόμηση:

$$\begin{array}{l} (x_1, x_2) \begin{cases} \rightarrow z_1 = w_{11}x_1 + w_{12}x_2 + w_{10} \\ \rightarrow z_2 = w_{21}x_1 + w_{22}x_2 + w_{20} \\ \rightarrow z_3 = w_{31}x_1 + w_{32}x_2 + w_{30} \end{cases} \end{array}$$

$$y_i = \frac{\exp(z_i)}{\exp(z_1) + \exp(z_2) + \exp(z_3)}$$

Κωνσταντίνος Καραμανής

```
class ThreeClassLogisticNN(nn.Module):
    def __init__(self):
        super(ThreeClassLogisticNN, self).__init__()
        self.linear = nn.Linear(2, 3)

    def forward(self, x):
        x = self.linear(x)
        x = nn.functional.softmax(x, dim=1)
        return x
```

# Λογιστική Παλινδρόμηση: Τι κάνει:

- Ο κώδικας κάνει η λογιστική παλινδρόμηση:

## Γεωμετρική έννοια αυτής της πράξης

$$\begin{aligned} z_1 &= w_{11}x_1 + w_{12}x_2 + w_{10} \\ z_2 &= w_{21}x_1 + w_{22}x_2 + w_{20} \\ z_3 &= w_{31}x_1 + w_{32}x_2 + w_{30} \end{aligned}$$

$$y_i = \frac{\exp(z_i)}{\exp(z_1) + \exp(z_2) + \exp(z_3)}$$

# Λογιστική Παλινδρόμηση: Τι κάνει:

- Ο κώδικας μας δείχνει ακριβώς τι κάνει η λογιστική παλινδρόμηση

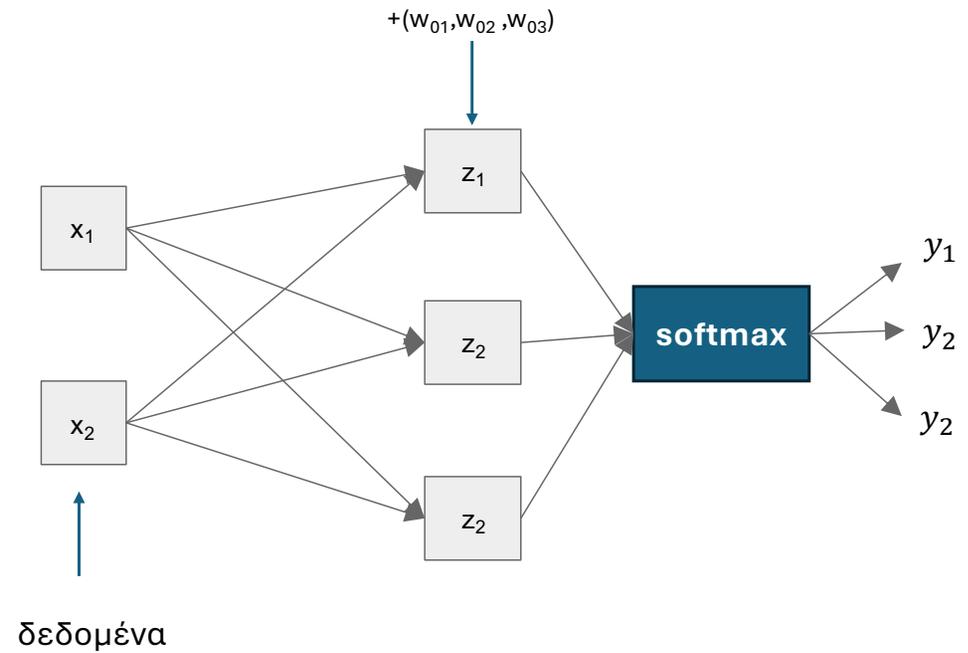
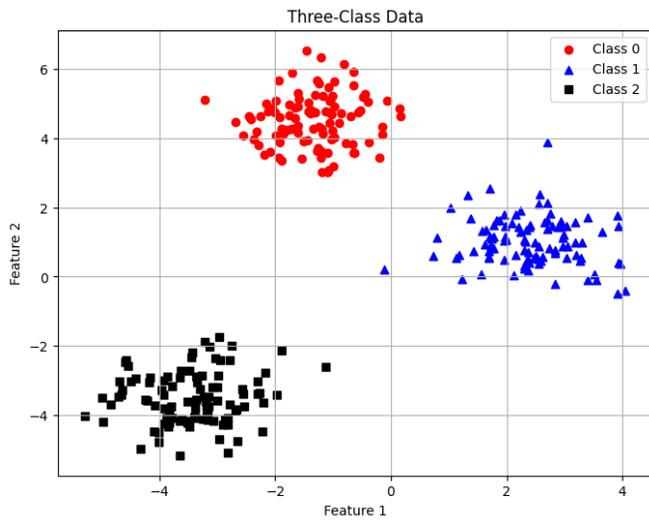
**Γεωμετρικά τι σημαίνουν αυτές οι παράμετροι;**

$(x_1, x_2)$

$$z_1 = W_{11}x_1 + W_{12}x_2 + W_{10}$$
$$z_2 = W_{21}x_1 + W_{22}x_2 + W_{20}$$
$$z_3 = W_{31}x_1 + W_{32}x_2 + W_{30}$$

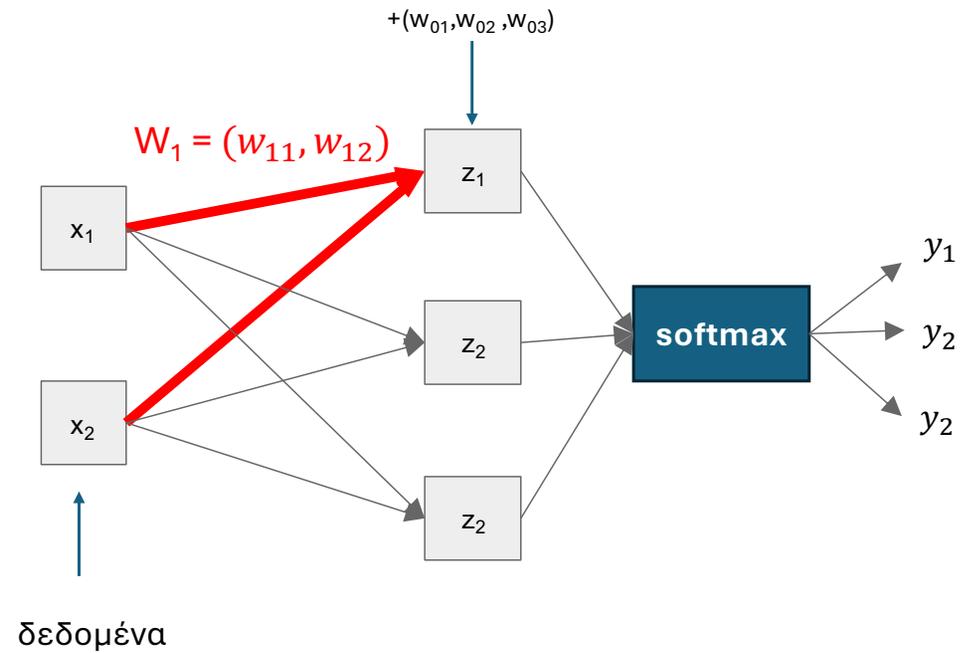
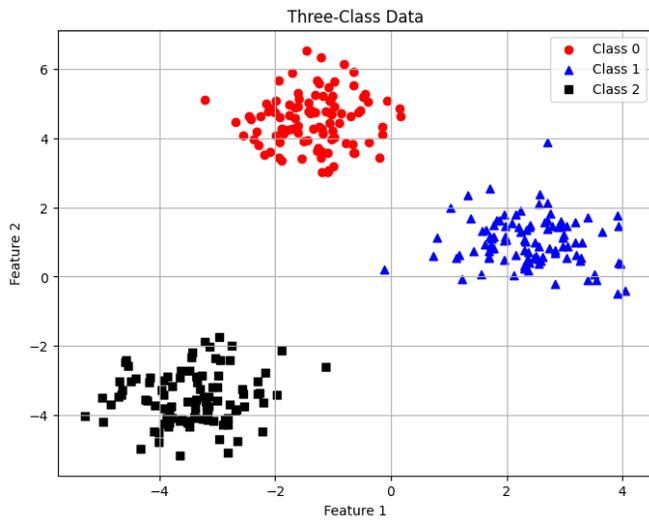
$$y_i = \frac{\exp(z_i)}{\exp(z_1) + \exp(z_2) + \exp(z_3)}$$

# Λογιστική Παλινδρόμηση: Τι κάνει;



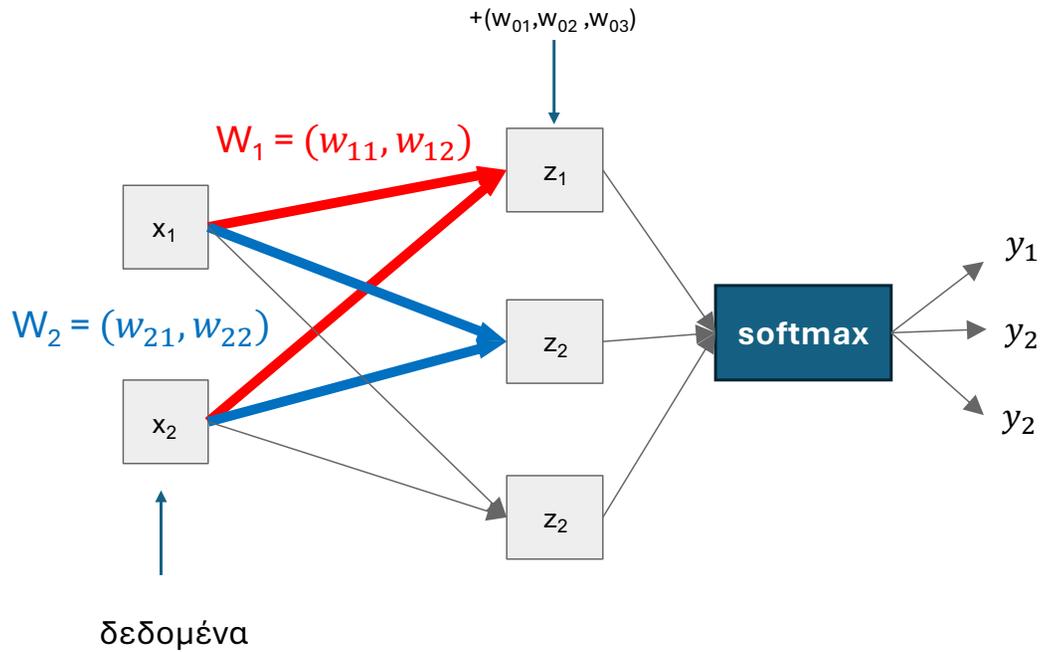
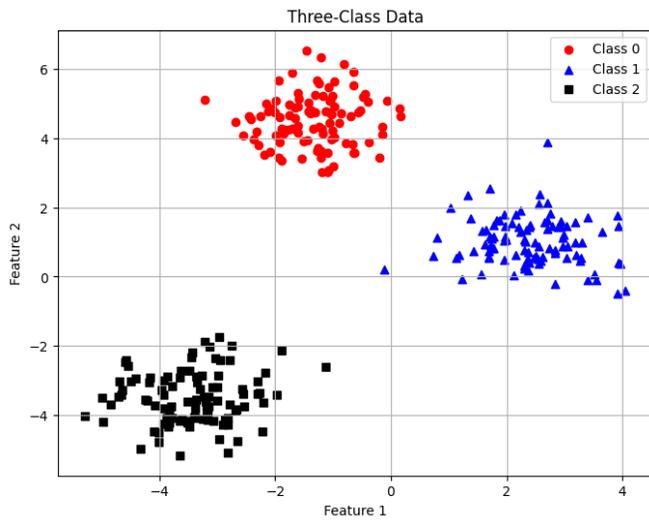
Κωνσταντίνος Καραμανής

# Λογιστική Παλινδρόμηση: Τι κάνει;

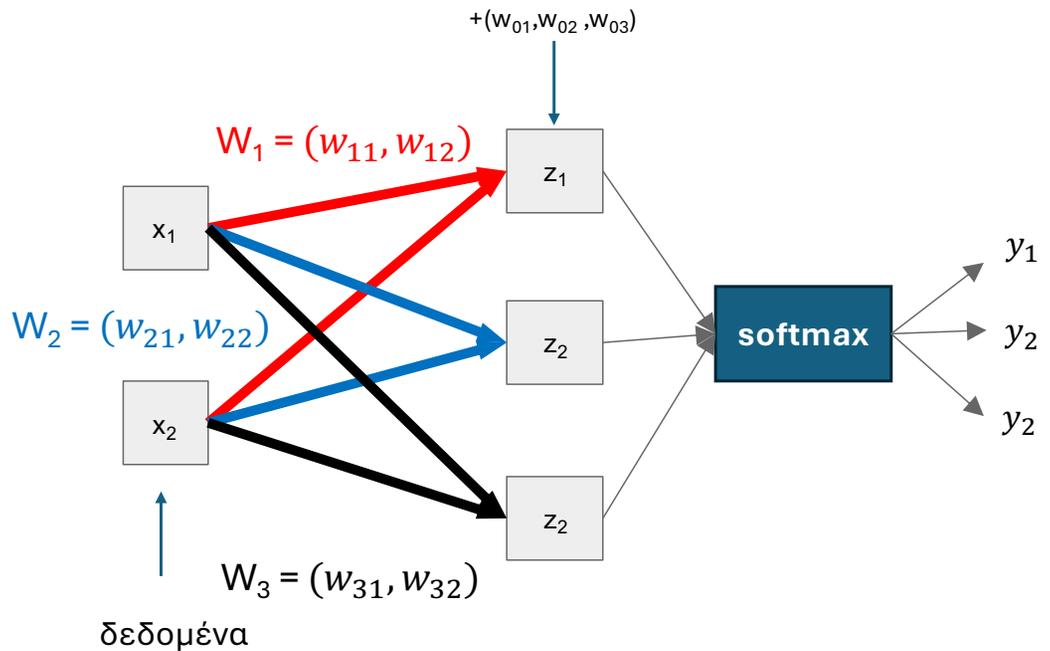
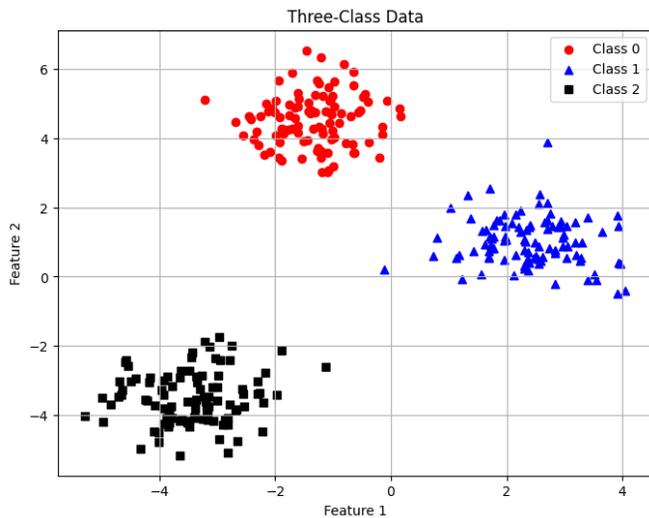


Κωνσταντίνος Καραμανής

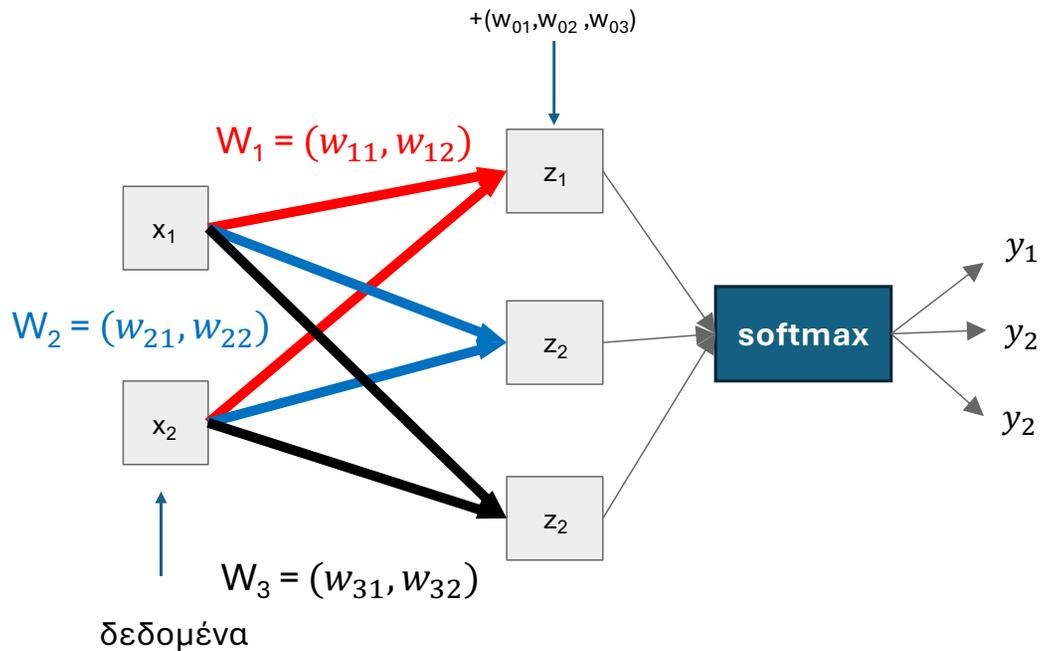
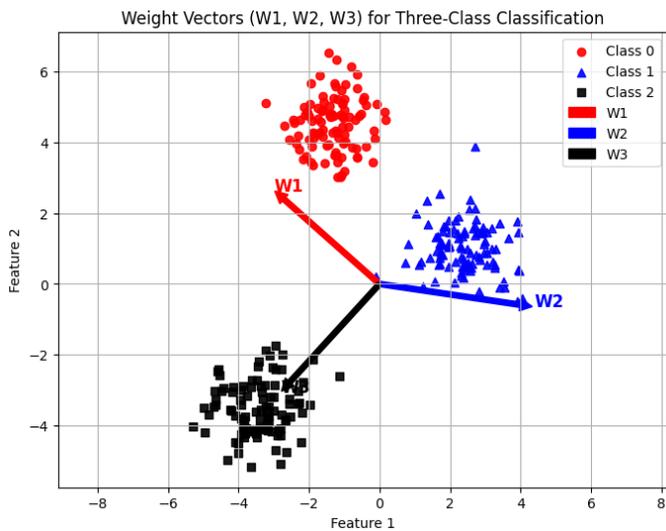
# Λογιστική Παλινδρόμηση: Τι κάνει;



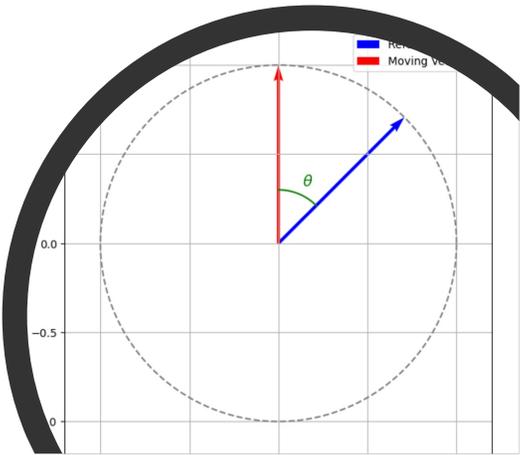
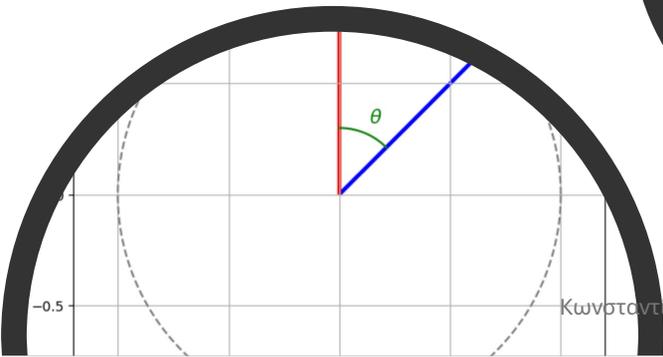
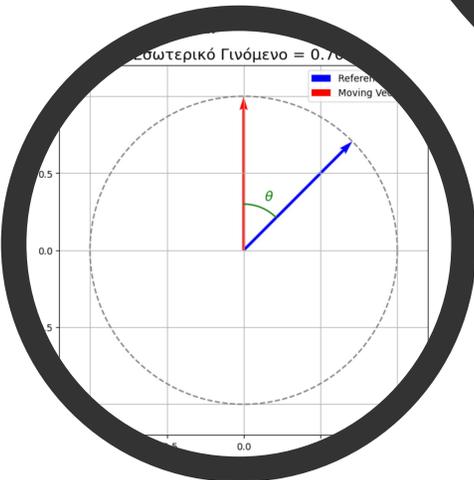
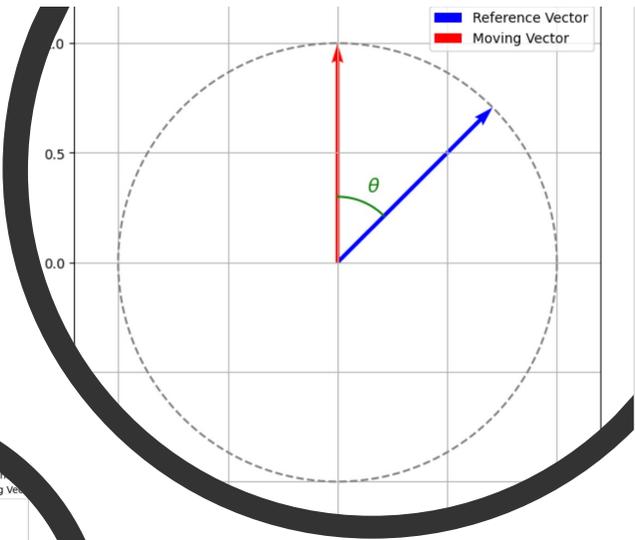
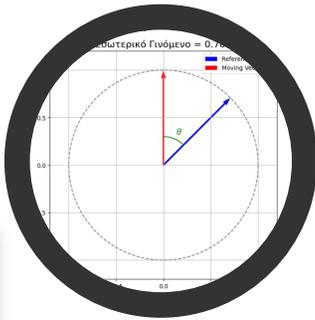
# Λογιστική Παλινδρόμηση: Τι κάνει;



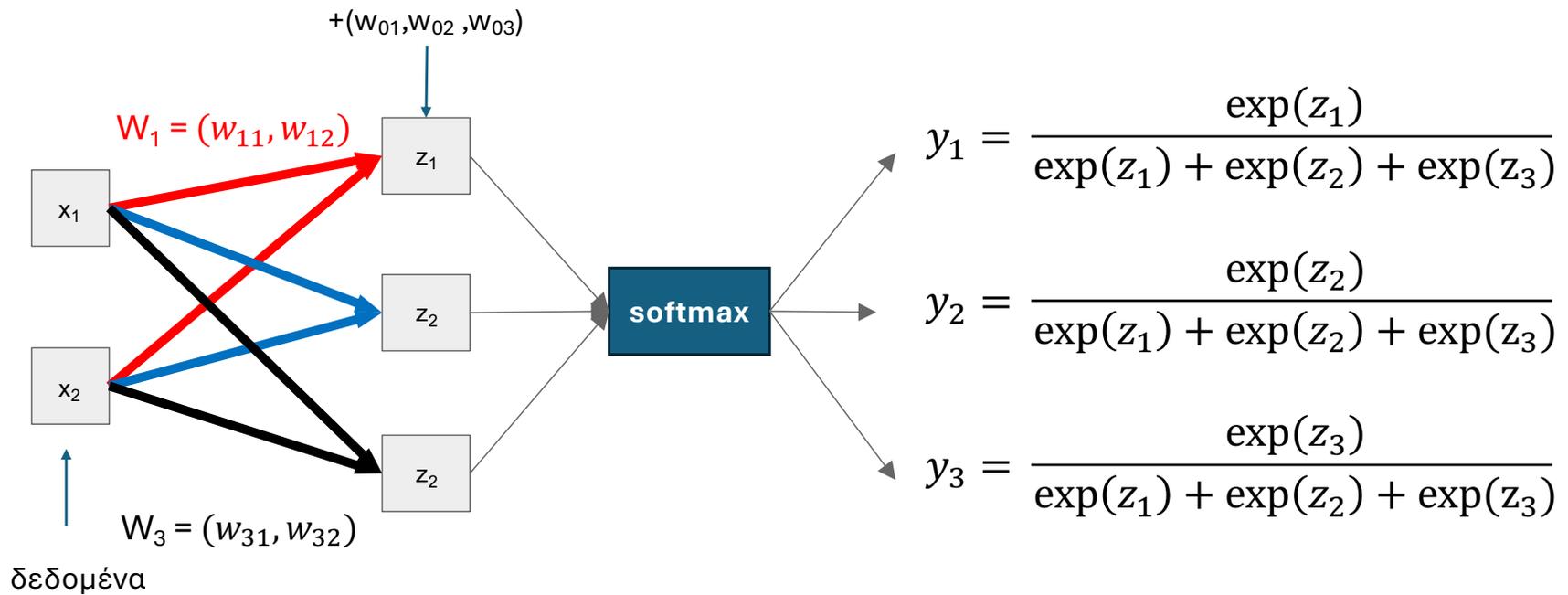
# Λογιστική Παλινδρόμηση: Τι κάνει;



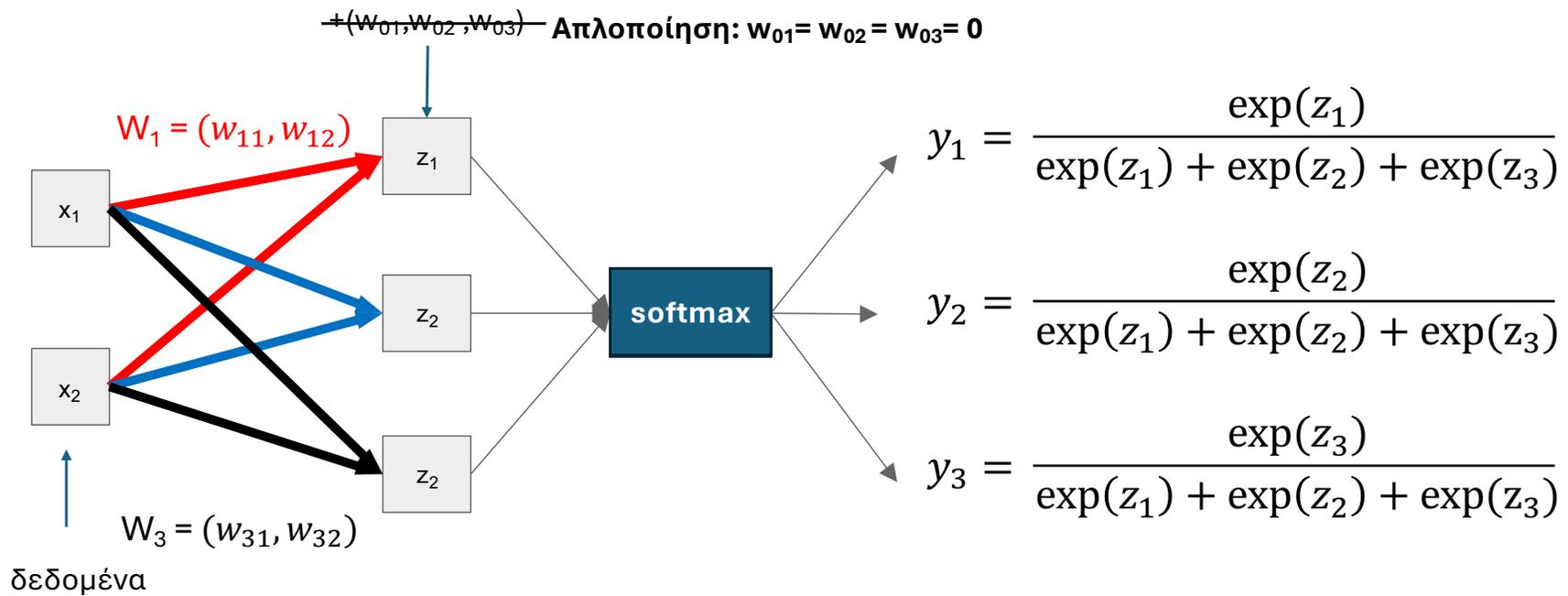
# Το Εσωτερικό Γινόμενο και η Γεωμετρία της Λογιστικής Παλινδρόμησης



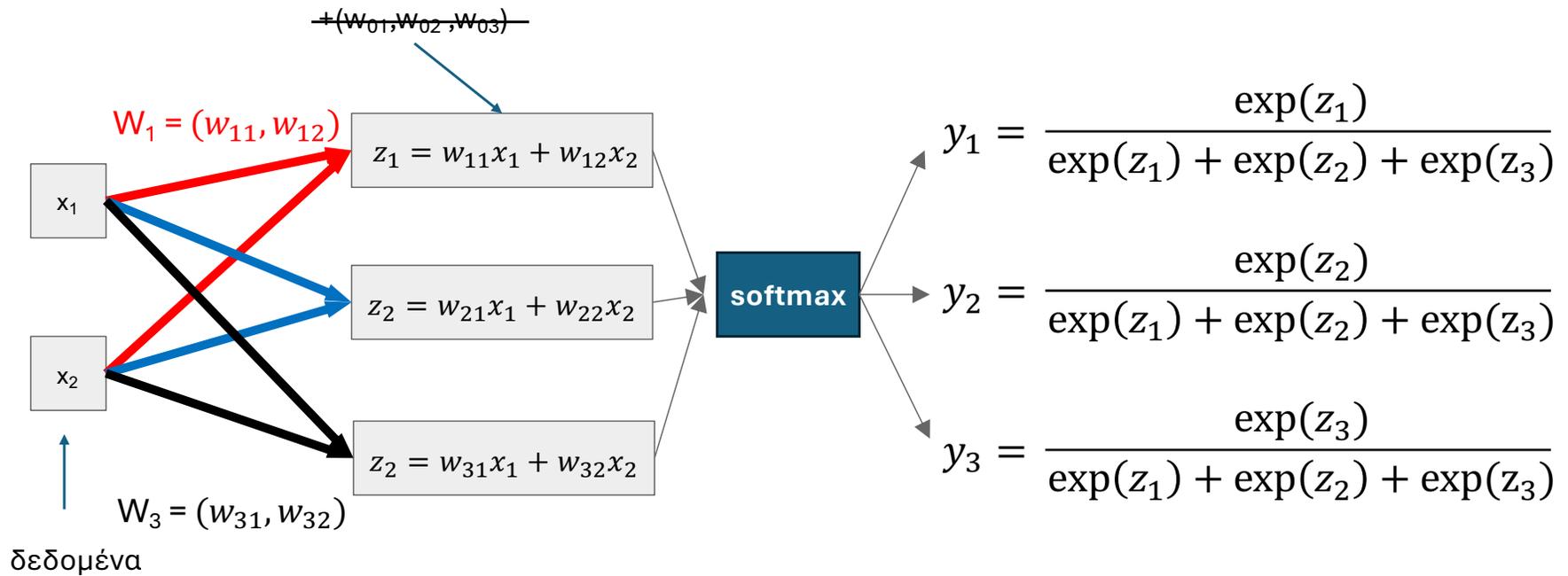
# Εσωτερικό Γινόμενο στη Λογιστική Παλινδρόμηση



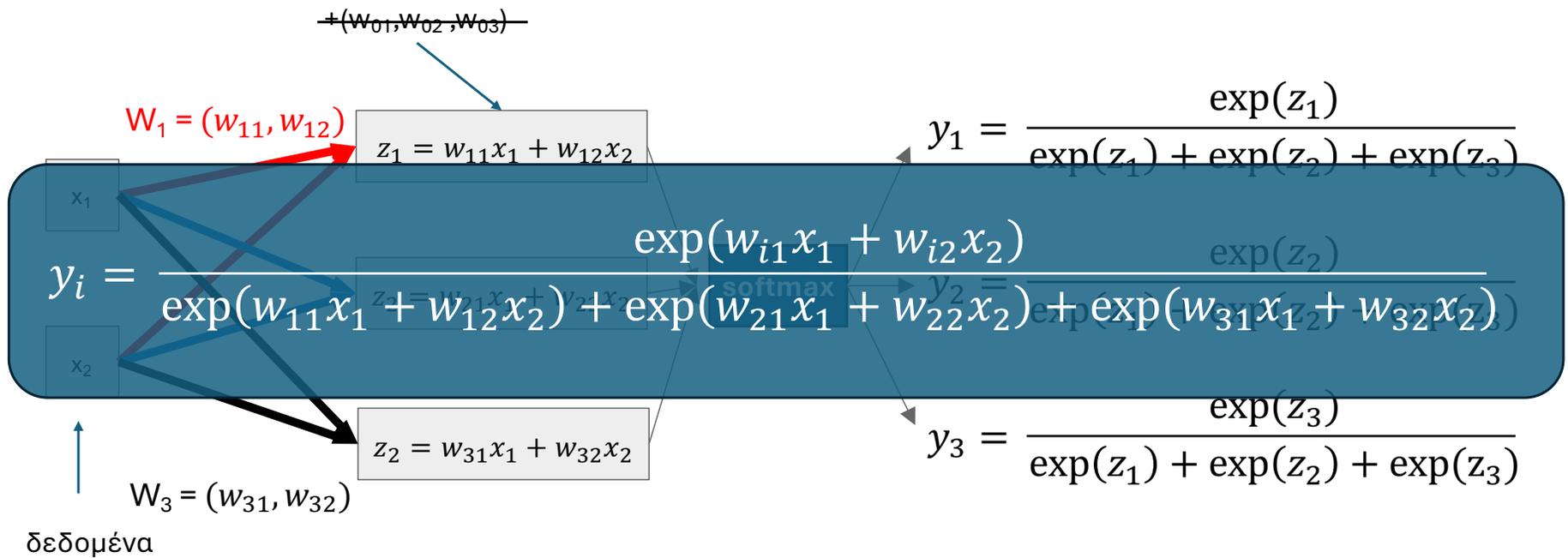
# Εσωτερικό Γινόμενο στη Λογιστική Παλινδρόμηση



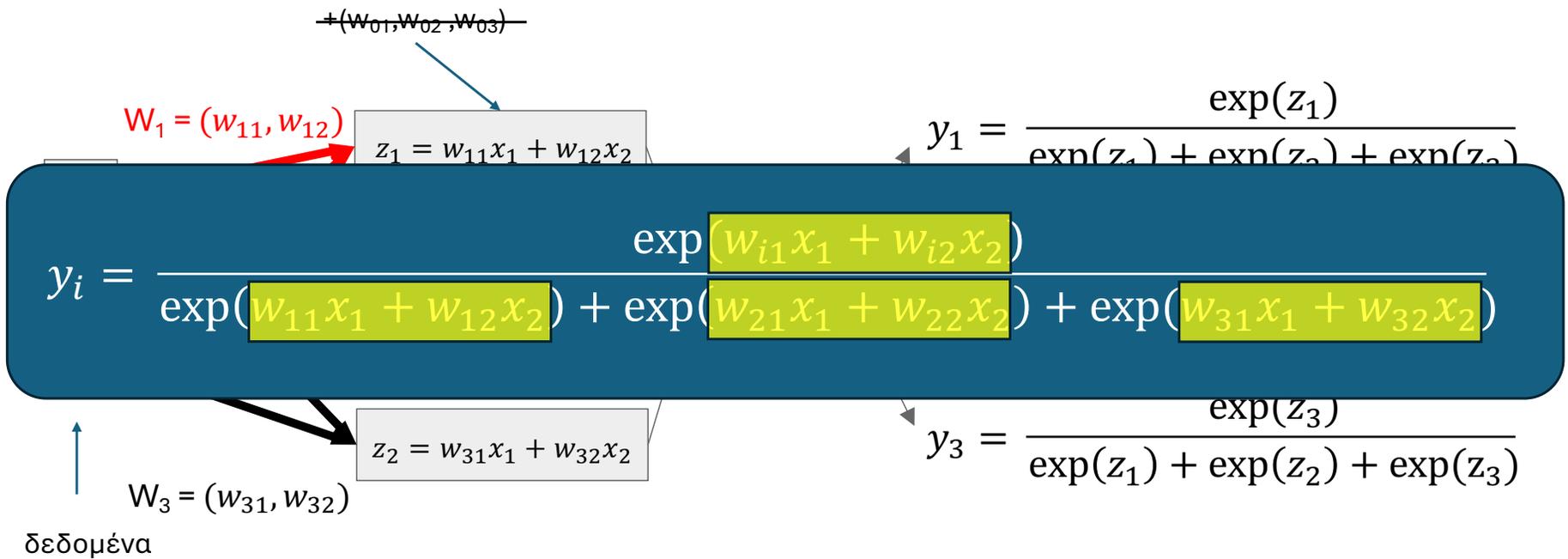
# Εσωτερικό Γινόμενο στη Λογιστική Παλινδρόμηση



# Εσωτερικό Γινόμενο στη Λογιστική Παλινδρόμηση



# Εσωτερικό Γινόμενο στη Λογιστική Παλινδρόμηση



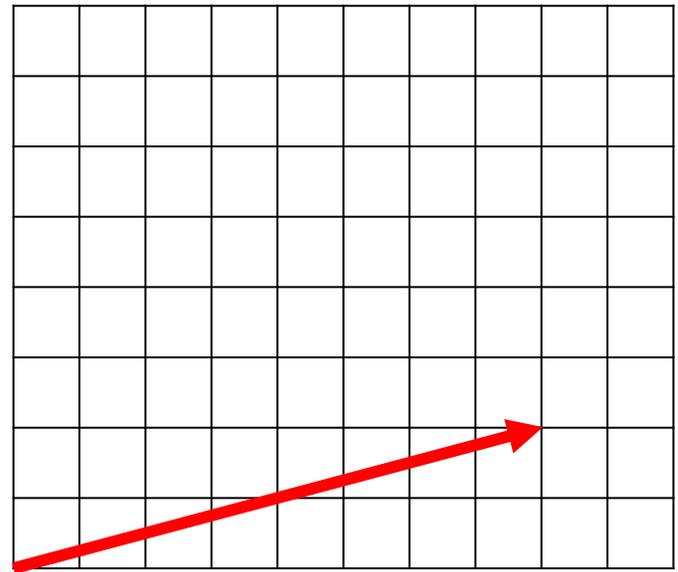
Εσωτερικό  
Γινόμενο

$$w_{11}x_1 + w_{12}x_2$$

Εσωτερικό  
Γινόμενο

$$w_{11}x_1 + w_{12}x_2$$

$w_1$

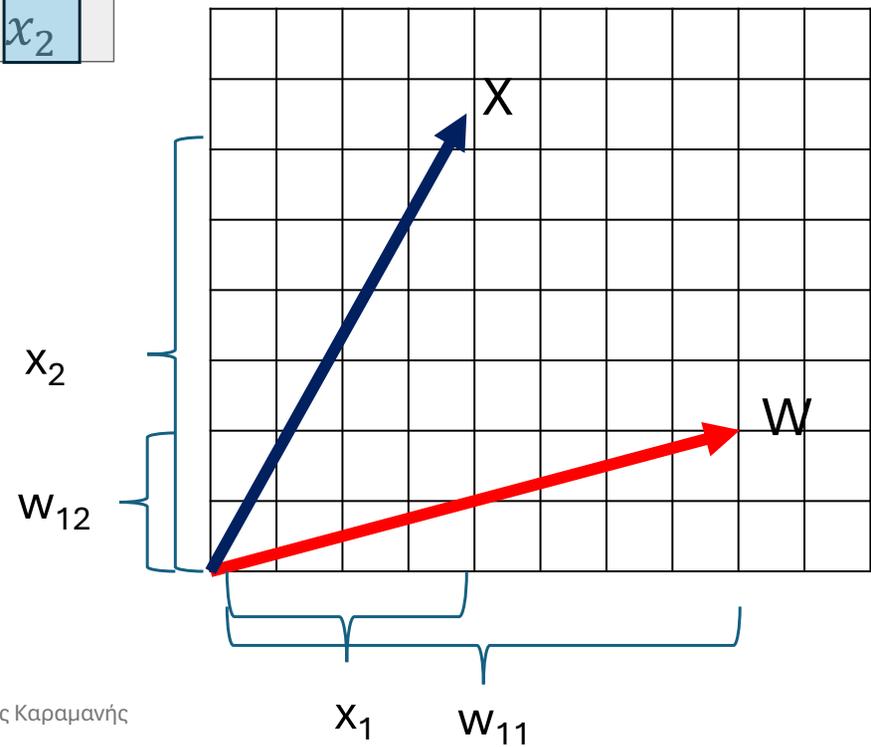


Κωνσταντίνος Καραμανής

$w_{11}$

# Εσωτερικό Γινόμενο

$$w_{11}x_1 + w_{12}x_2$$



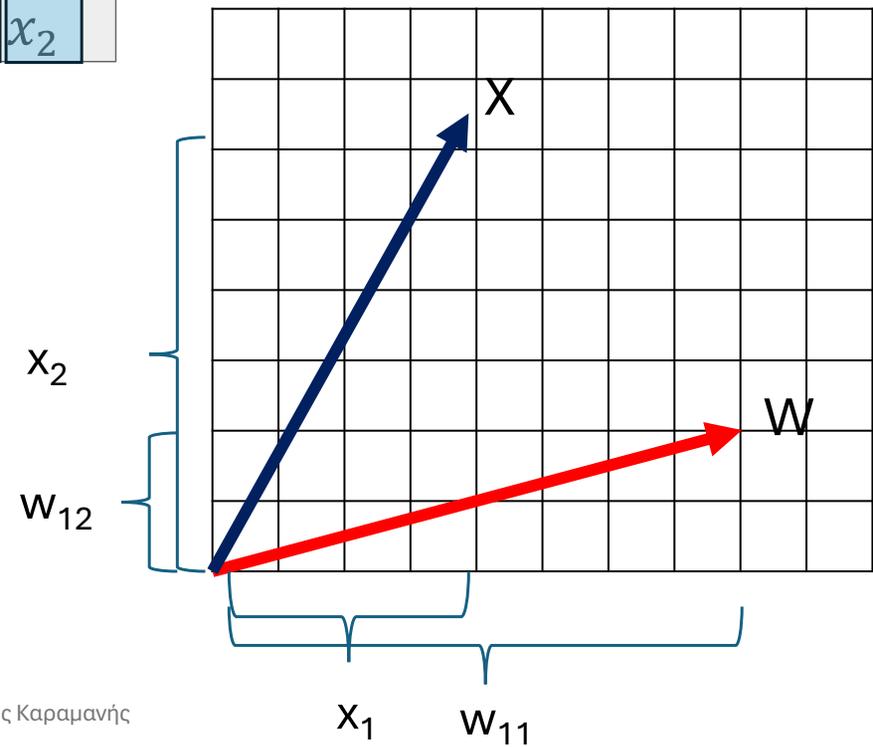
Κωνσταντίνος Καραμανής

# Εσωτερικό Γινόμενο

$$w_{11}x_1 + w_{12}x_2$$

$$W = (w_{11}, w_{12})$$

$$X = (x_1, x_2)$$



# Εσωτερικό Γινόμενο

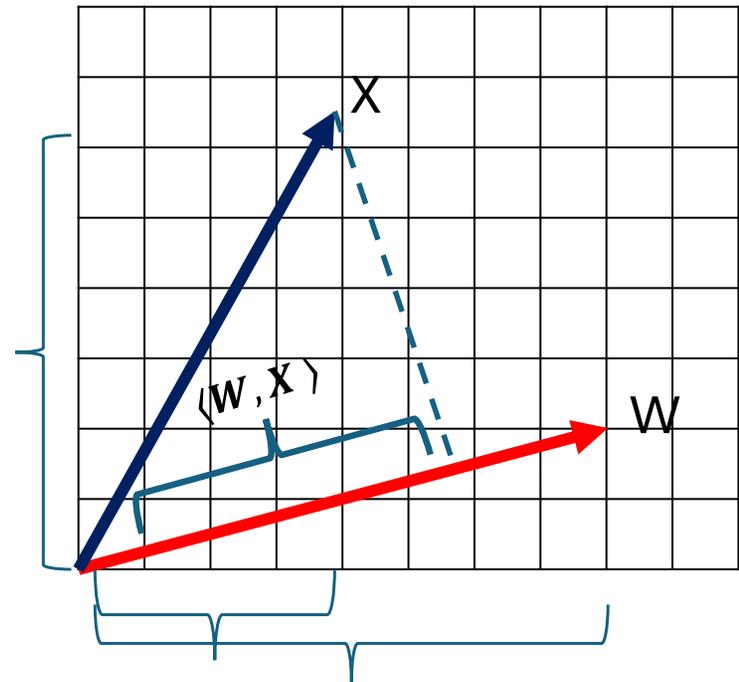
$$w_{11}x_1 + w_{12}x_2$$

$$W = (w_{11}, w_{12})$$

$$X = (x_1, x_2)$$

$x_2$

$w_{12}$



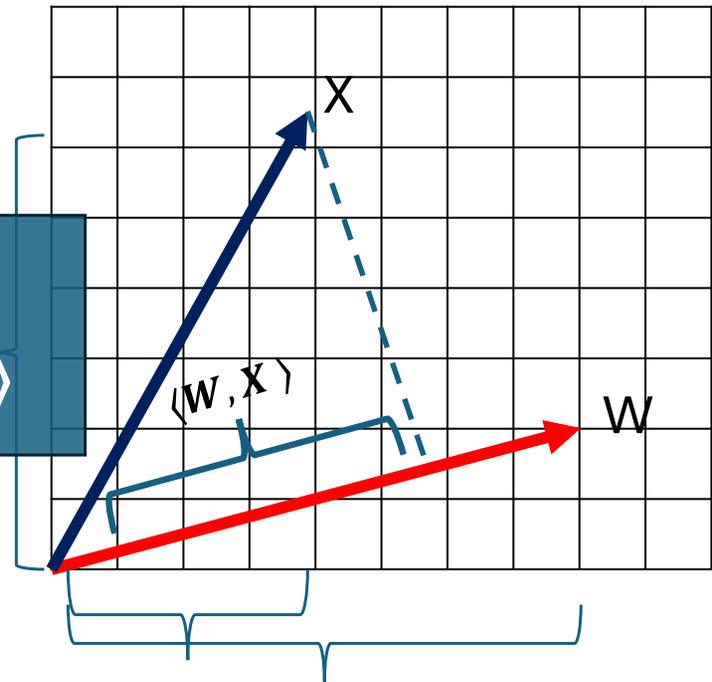
Εσωτερικό  
Γινόμενο

$$w_{11}x_1 + w_{12}x_2$$

$$W = (w_{11}, w_{12})$$

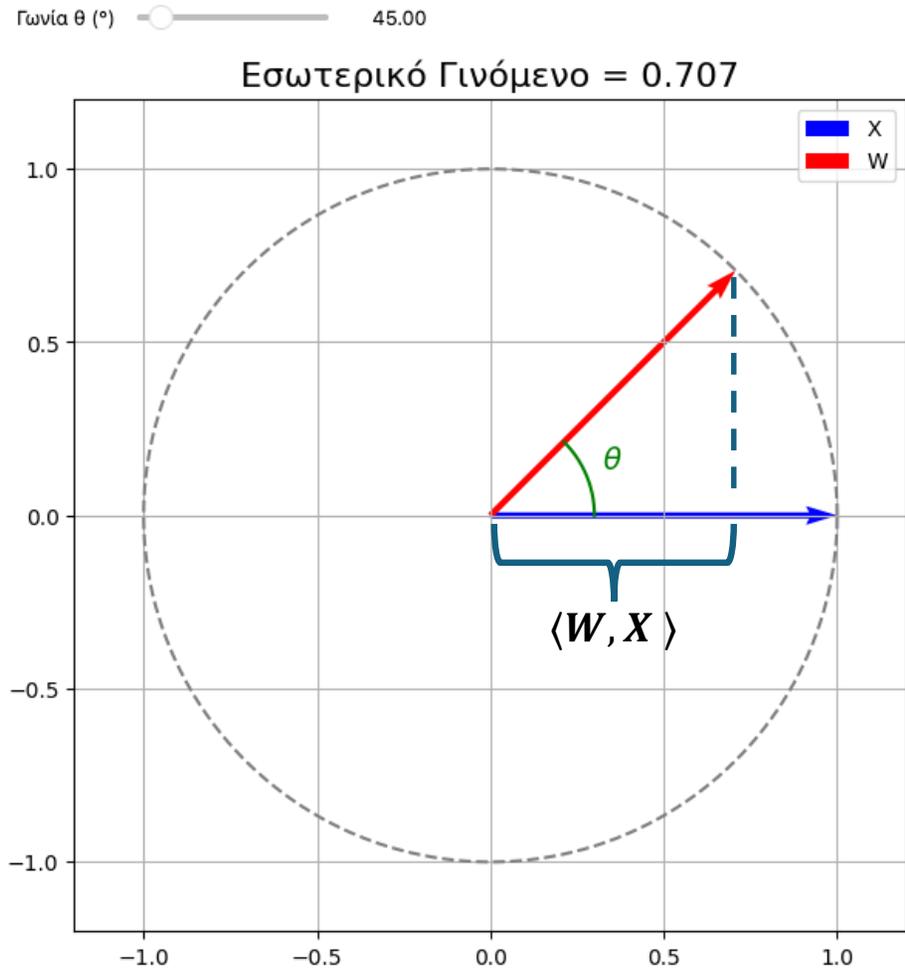
$$\text{Εσωτερικό Γινόμενο}$$
$$w_{11}x_1 + w_{12}x_2 = \langle W, X \rangle$$

$w_{12}$



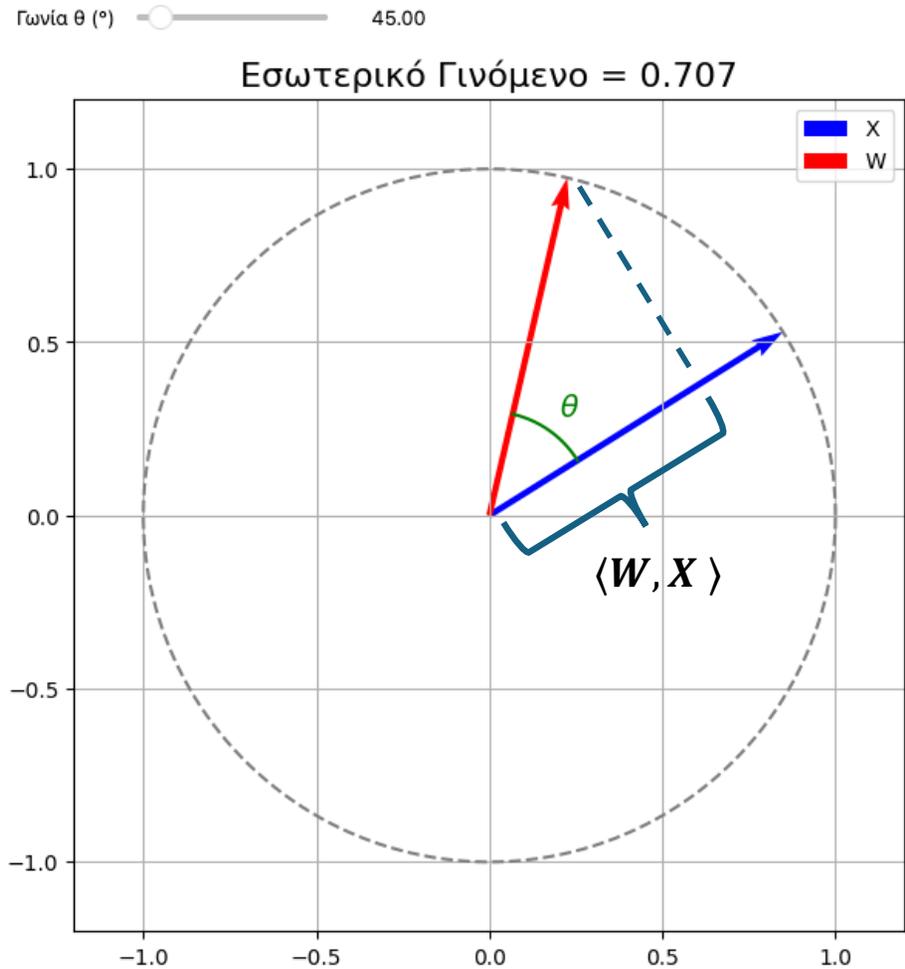
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



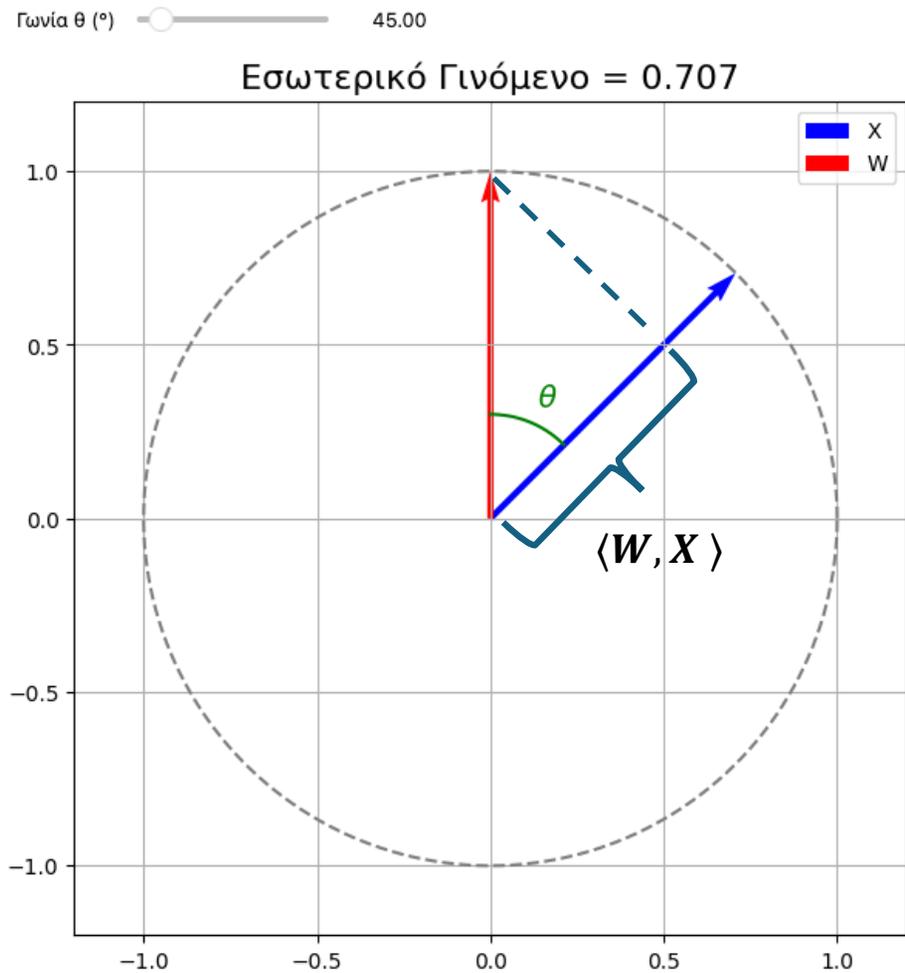
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



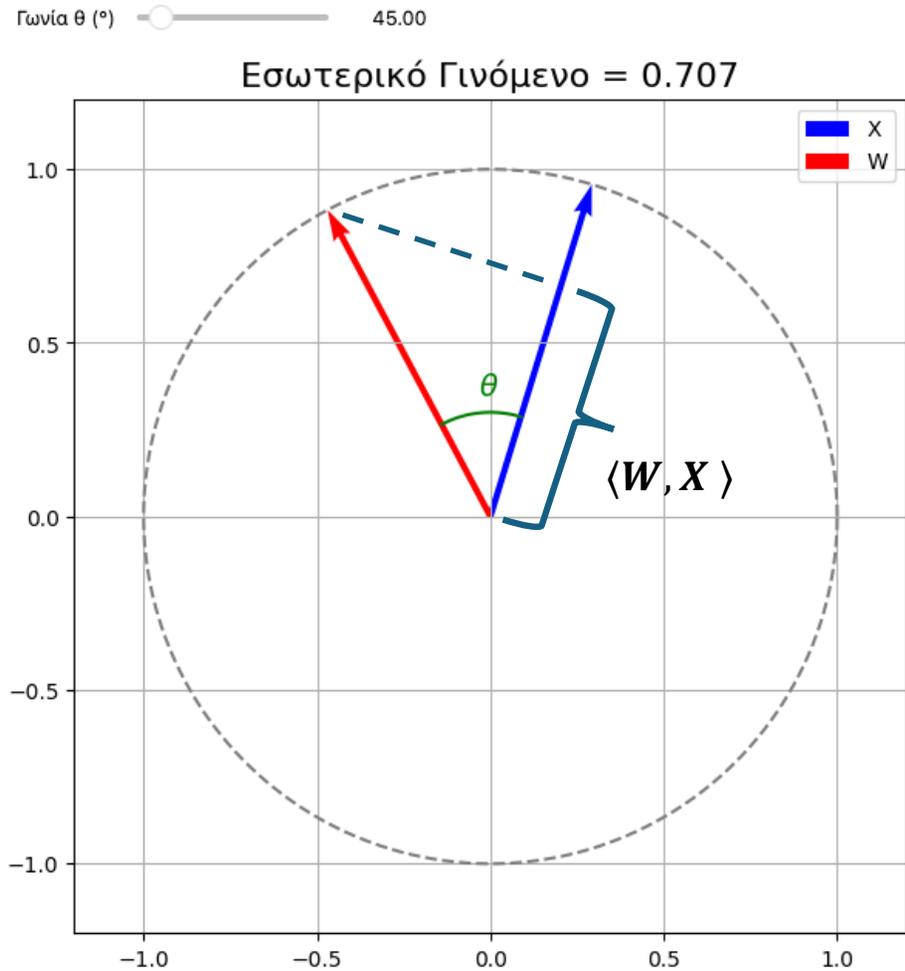
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



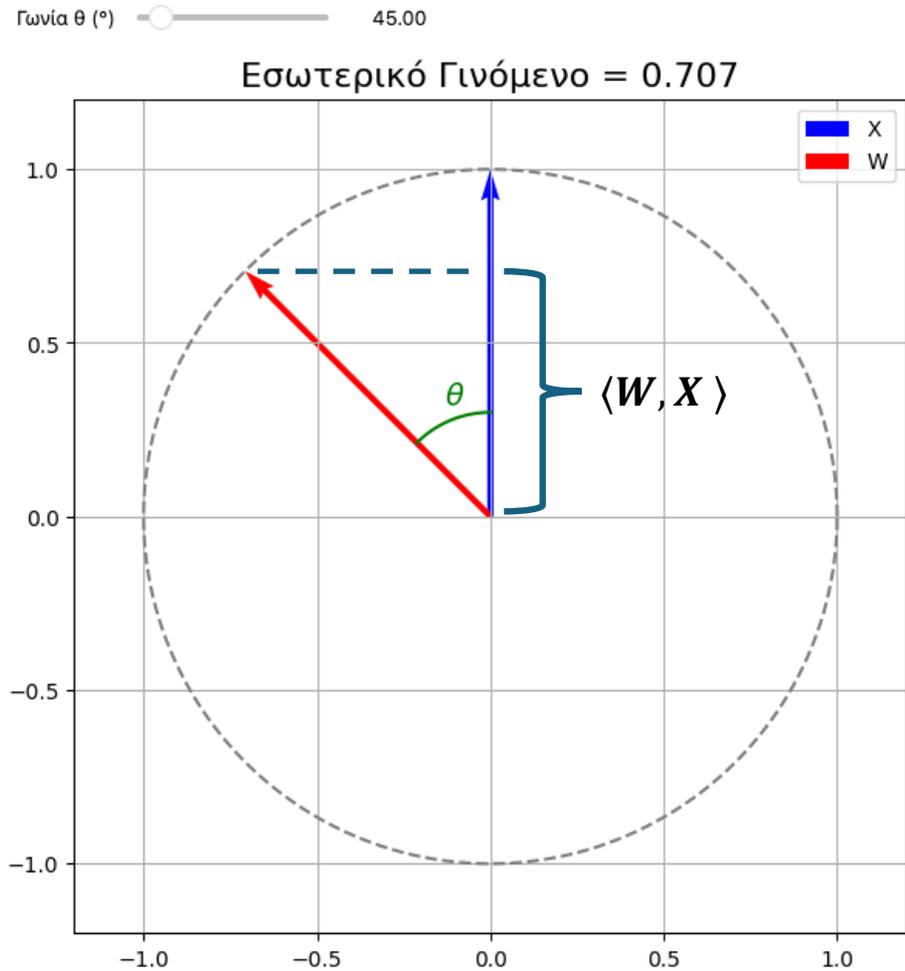
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



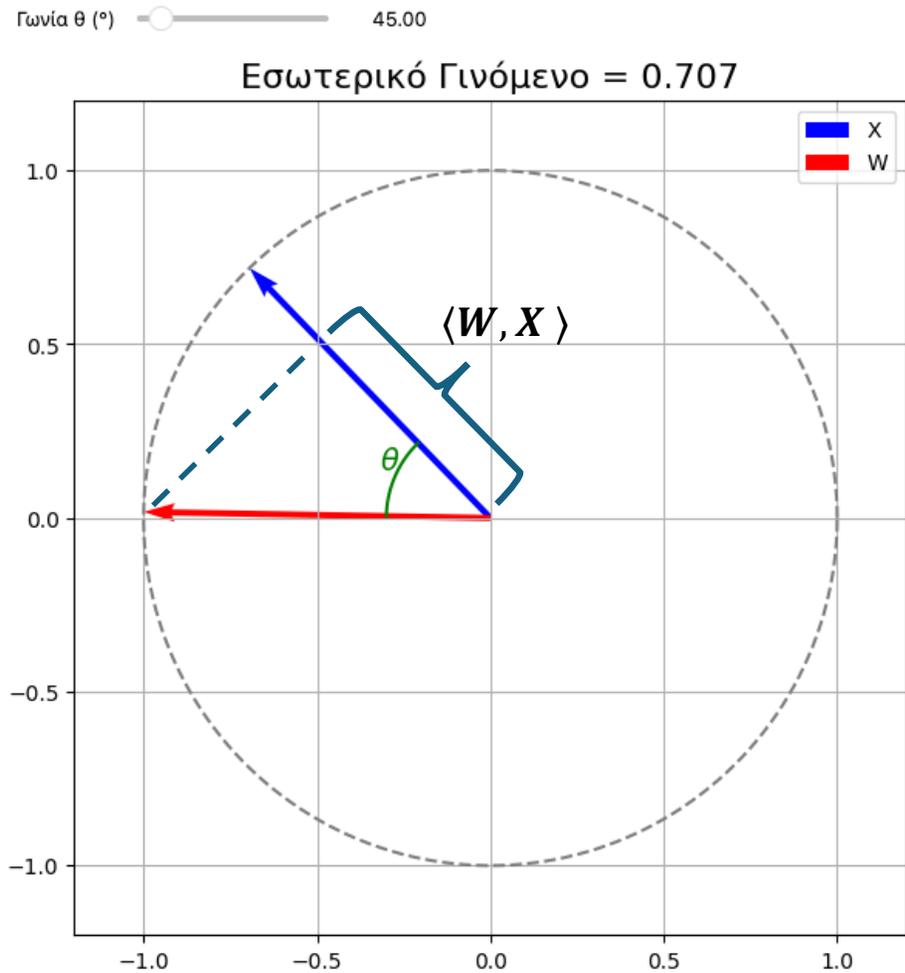
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



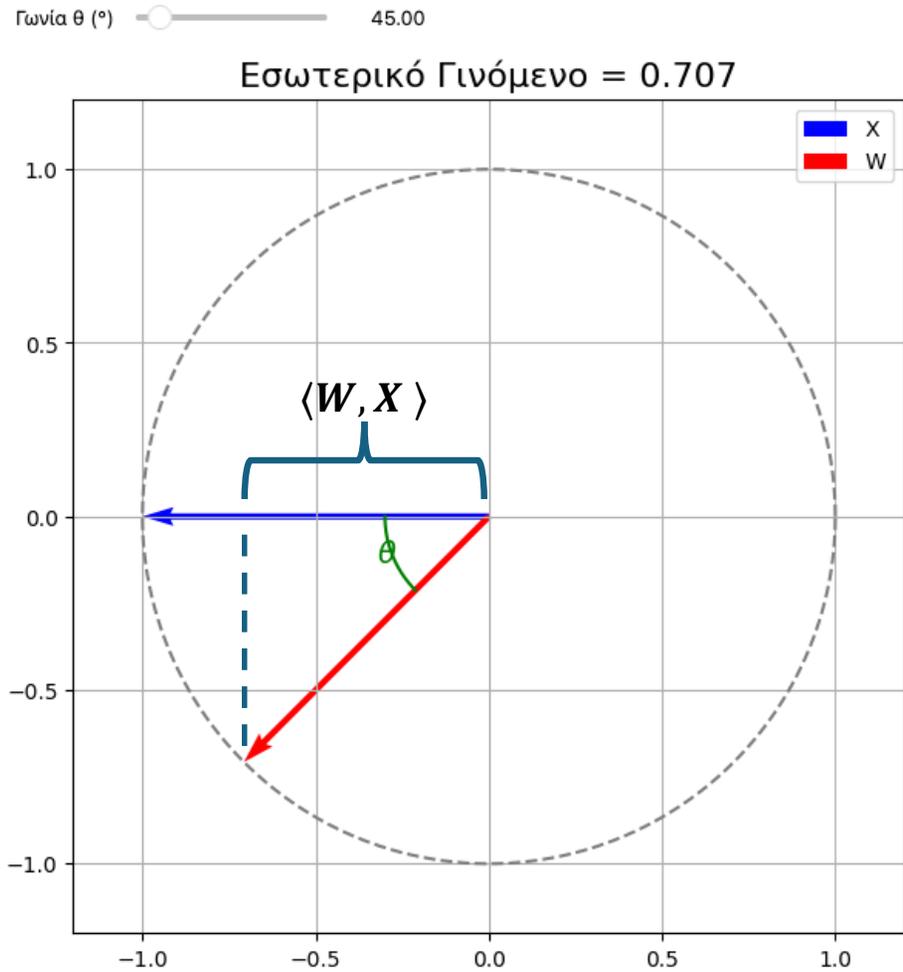
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



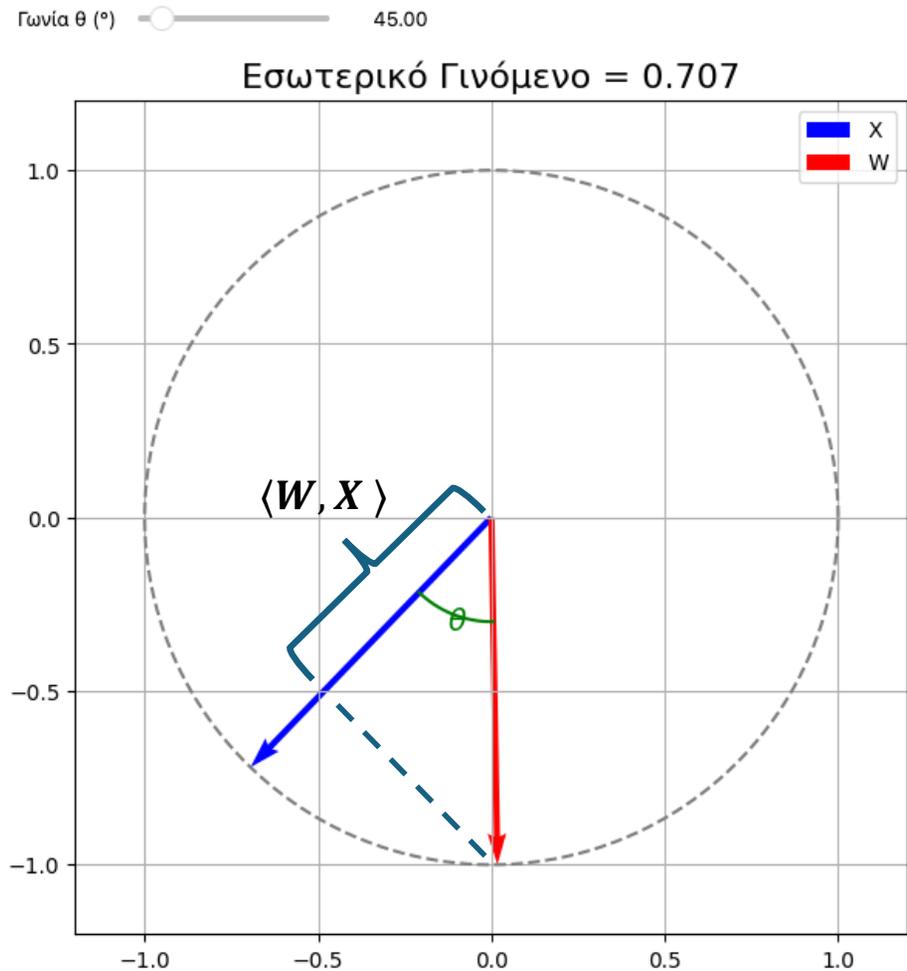
# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$



# Εσωτερικό Γινόμενο και η γωνία $\theta$

Για οποιαδήποτε δύο  
διανύσματα πάνω στον κύκλο,  
το εσωτερικό γινόμενο  
εξαρτάται μόνο από τη γωνία  $\theta$



# Εσωτερικό Γινόμενο και η γωνία $\theta$



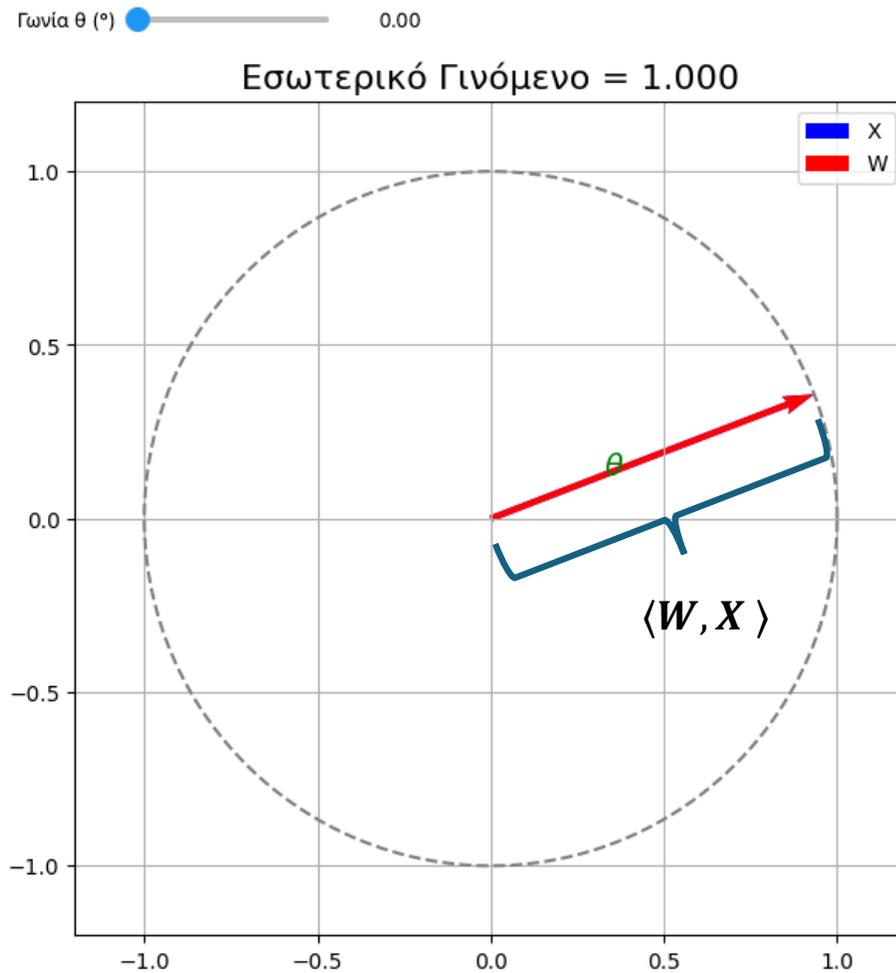
**Εσωτερικό Γινόμενο:  $\langle W, X \rangle$**   
Εξαρτάται μόνο από τη γωνία  $\theta$  μεταξύ των δύο διανυσμάτων πάνω στον κύκλο

Για οποιαδήποτε δυο διανύσματα πάνω στον κύκλο, το εσωτερικό γινόμενο εξαρτάται μόνο από τη γωνία  $\theta$

# Εσωτερικό Γινόμενο και η γωνία $\theta$

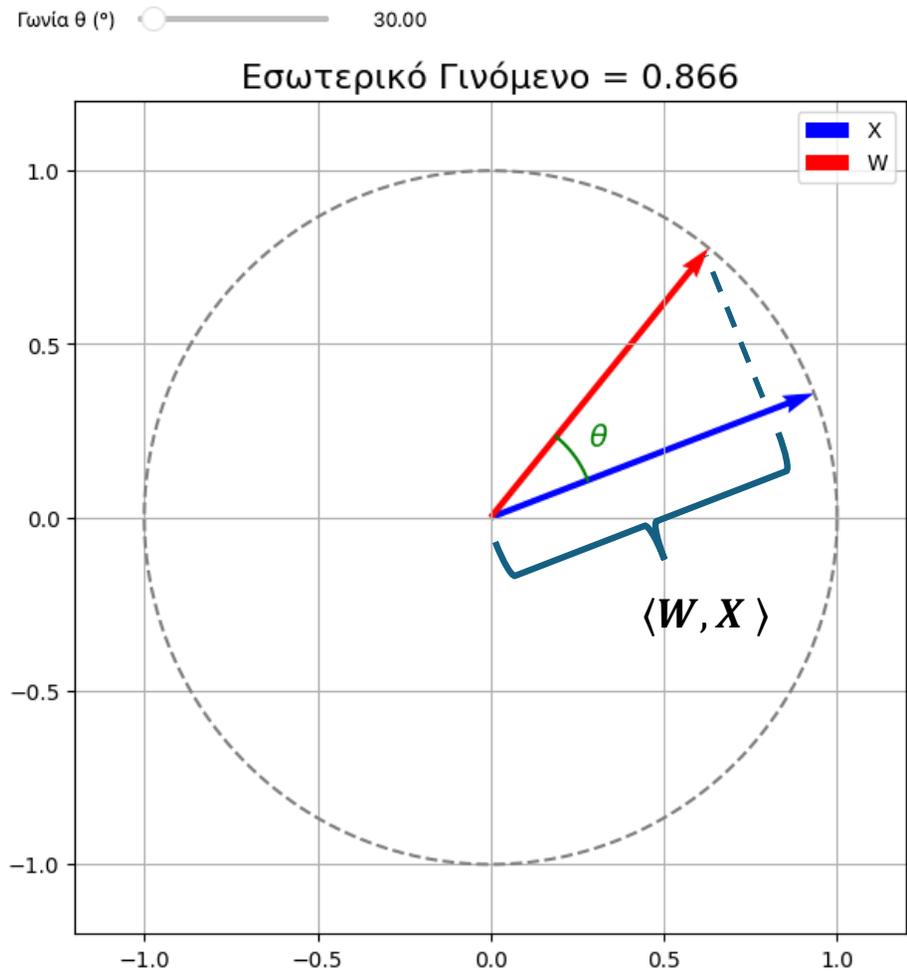
- $\theta = 0^\circ: \langle W, X \rangle = 1$

- 
- 
- 
- 
- 
- 
- 
- 



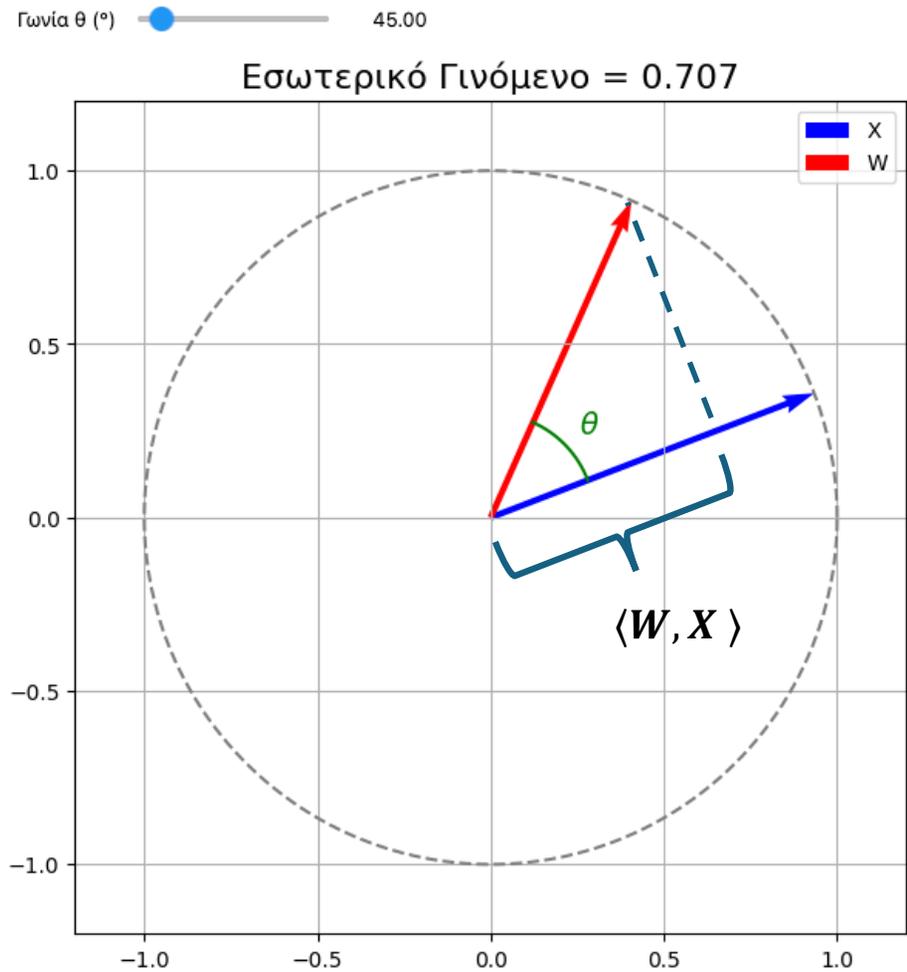
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ$ :  $\langle W, X \rangle = 1$
- $\theta = 30^\circ$ :  $\langle W, X \rangle = 0.866$
- 
- 
- 
- 
- 
- 
- 



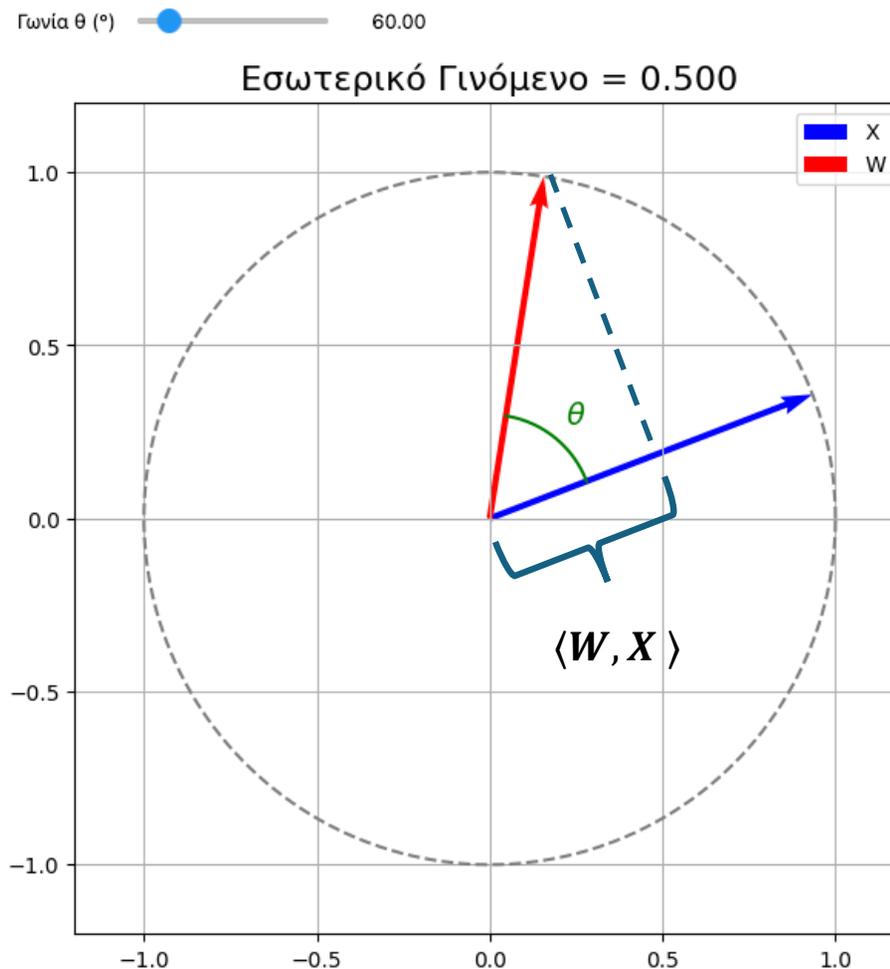
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- 
- 
- 
- 
- 
- 



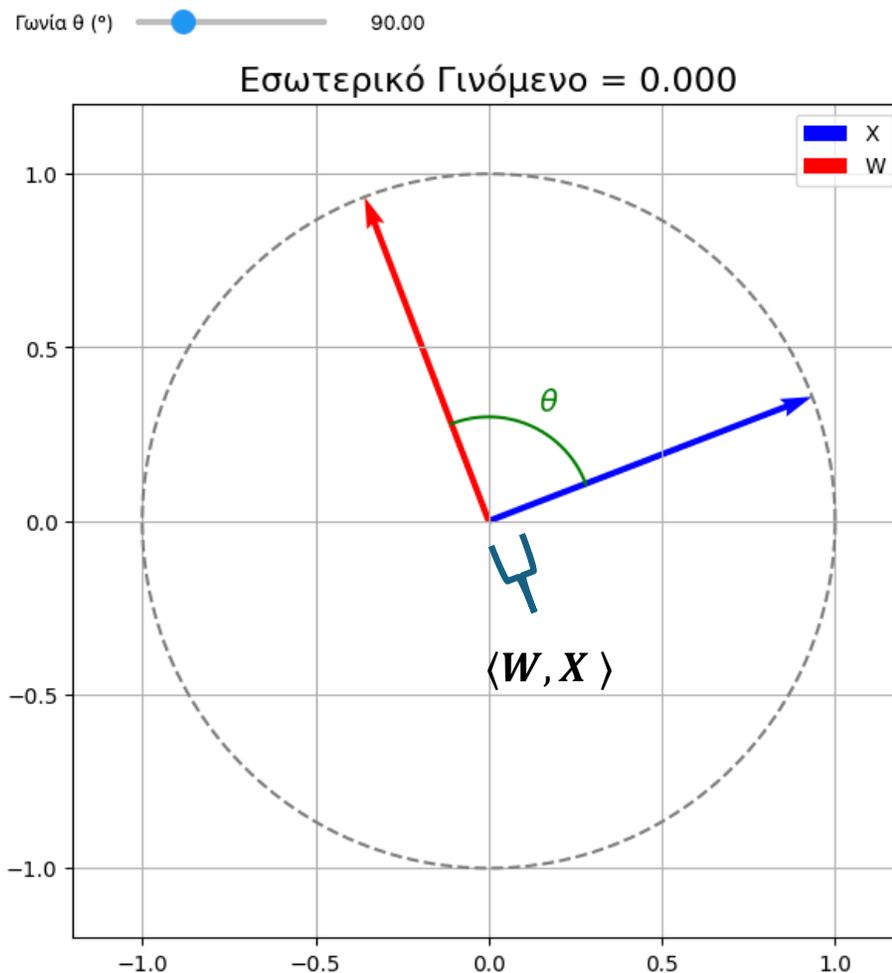
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- 
- 
- 
- 
- 



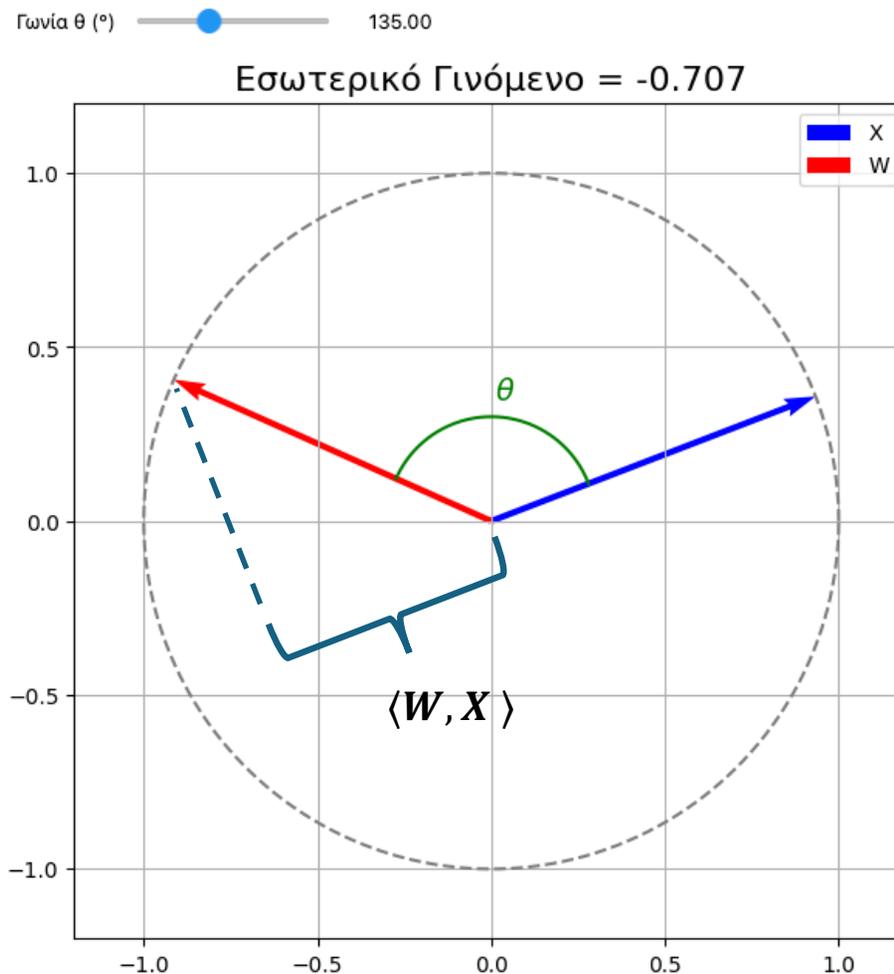
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- $\theta = 90^\circ: \langle W, X \rangle = 0$
- 
- 
- 
- 



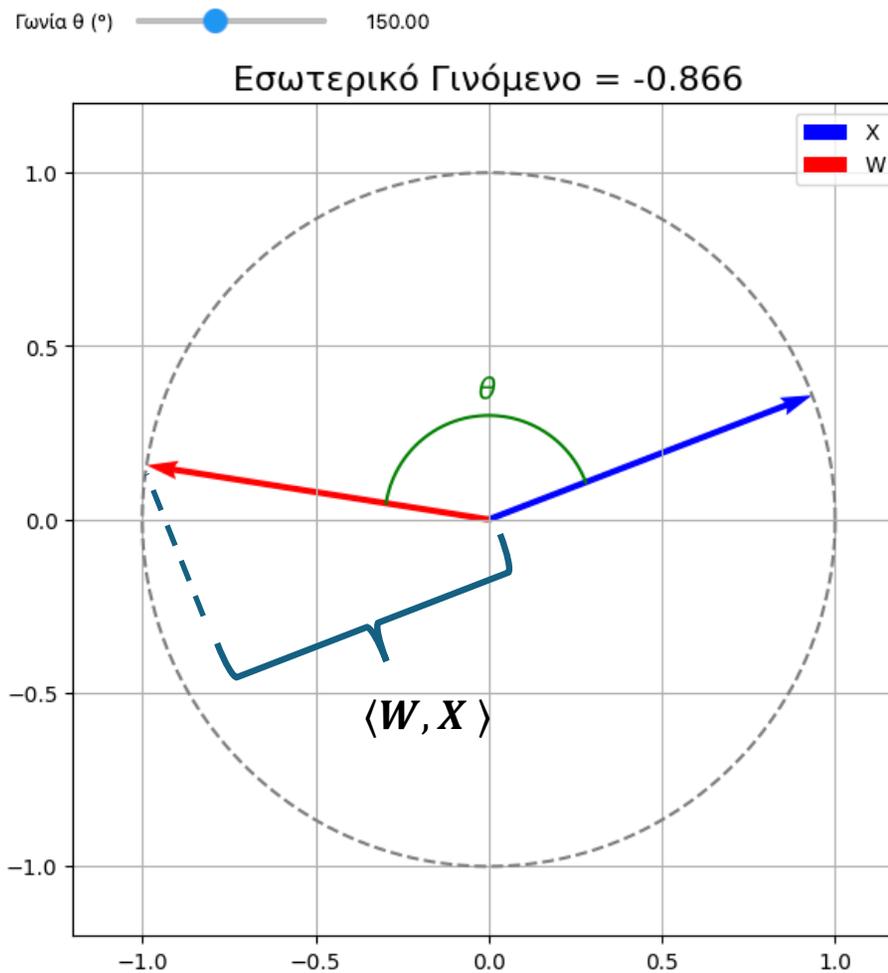
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- $\theta = 90^\circ: \langle W, X \rangle = 0$
- $\theta = 135^\circ: \langle W, X \rangle = -0.707$
- 
- 
- 



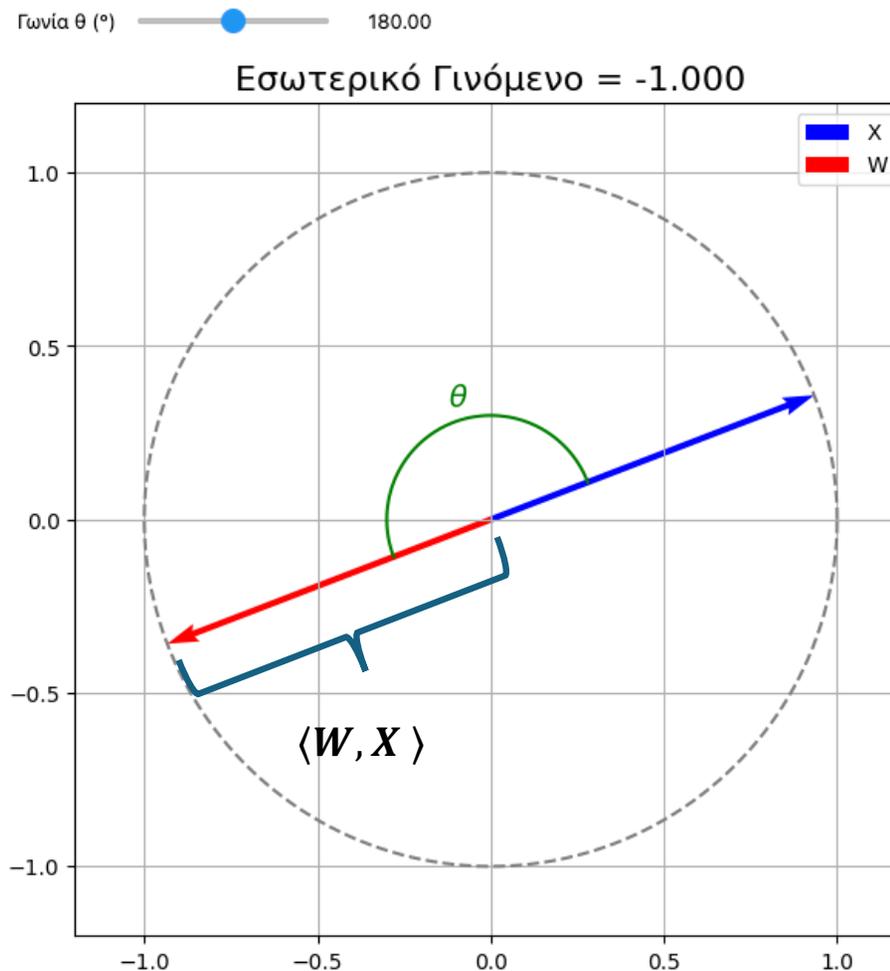
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- $\theta = 90^\circ: \langle W, X \rangle = 0$
- $\theta = 135^\circ: \langle W, X \rangle = -0.707$
- $\theta = 150^\circ: \langle W, X \rangle = -0.866$
- 
- 



# Εσωτερικό Γινόμενο και η γωνία $\theta$

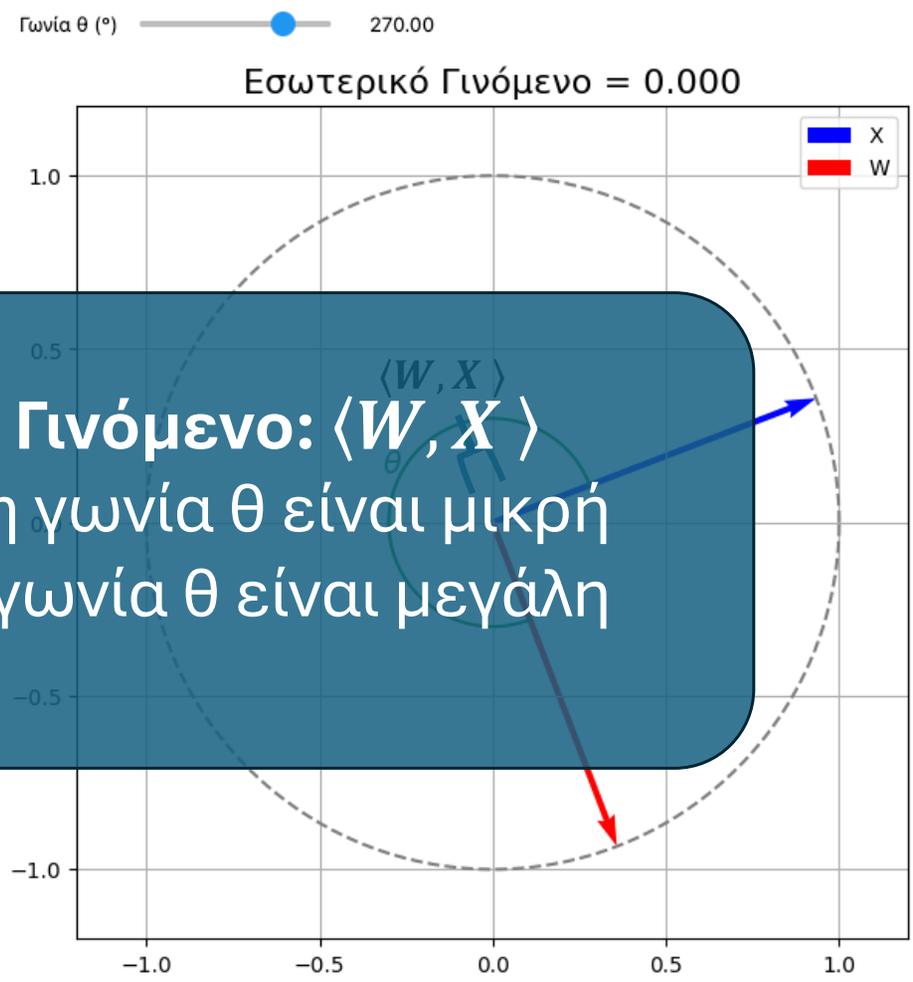
- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- $\theta = 90^\circ: \langle W, X \rangle = 0$
- $\theta = 135^\circ: \langle W, X \rangle = -0.707$
- $\theta = 150^\circ: \langle W, X \rangle = -0.866$
- $\theta = 180^\circ: \langle W, X \rangle = -1$
- 



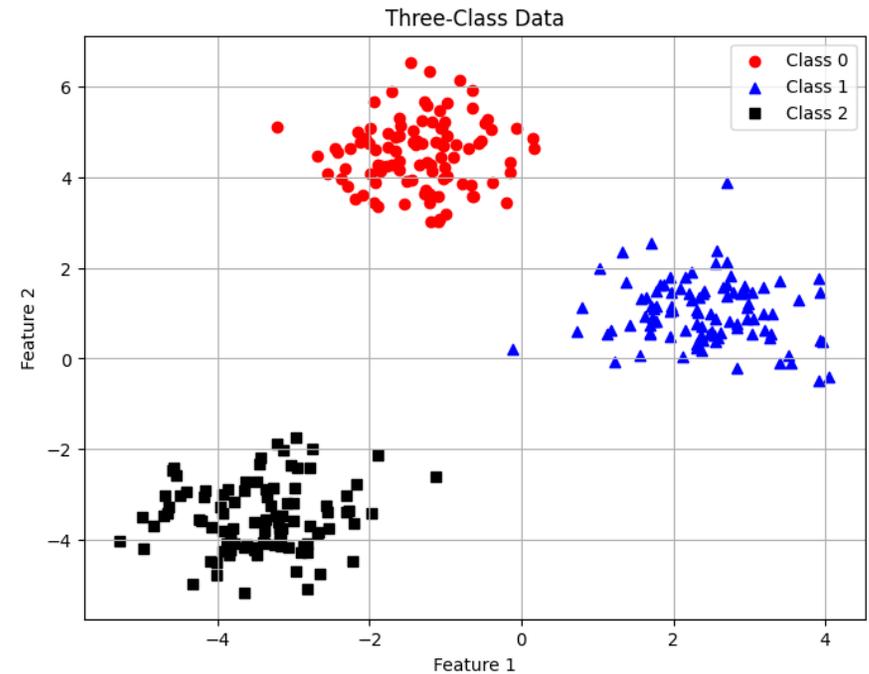
# Εσωτερικό Γινόμενο και η γωνία $\theta$

- $\theta = 0^\circ: \langle W, X \rangle = 1$
- $\theta = 30^\circ: \langle W, X \rangle = 0.866$
- $\theta = 45^\circ: \langle W, X \rangle = 0.707$
- $\theta = 60^\circ: \langle W, X \rangle = 0.5$
- $\theta = 90^\circ: \langle W, X \rangle = 0$
- $\theta = 135^\circ: \langle W, X \rangle = -0.707$
- $\theta = 150^\circ: \langle W, X \rangle = -0.866$
- $\theta = 180^\circ: \langle W, X \rangle = -1$
- $\theta = 270^\circ: \langle W, X \rangle = 0$

**Εσωτερικό Γινόμενο:  $\langle W, X \rangle$**   
Μεγάλο όταν η γωνία  $\theta$  είναι μικρή  
Μικρό όταν η γωνία  $\theta$  είναι μεγάλη



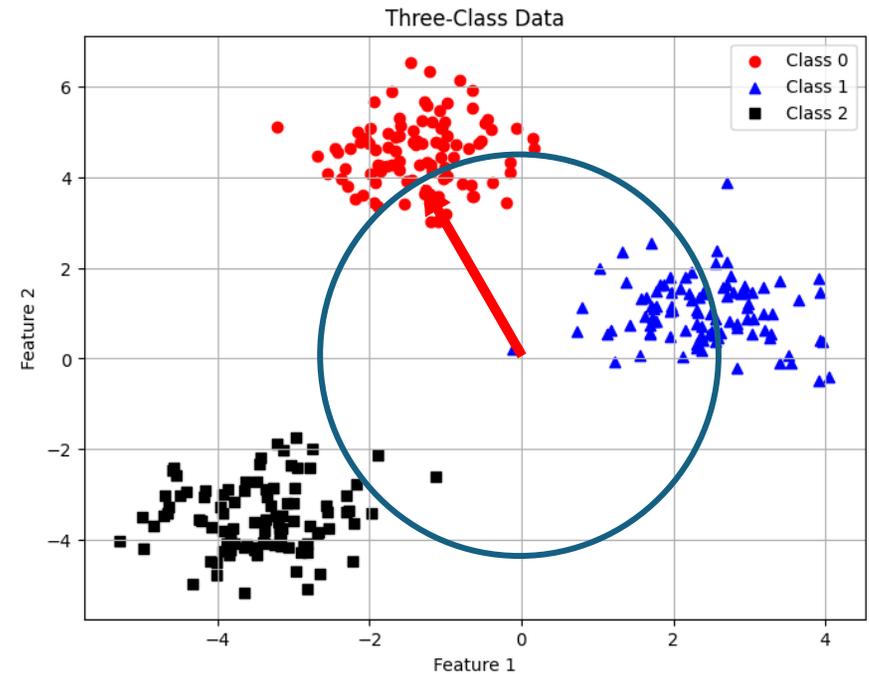
Πού δείχνει το  $W_1 = (w_{11}, w_{12})$



$$y_1 = \frac{\exp(w_{11}x_1 + w_{12}x_2)}{\exp(w_{11}x_1 + w_{12}x_2) + \exp(w_{21}x_1 + w_{22}x_2) + \exp(w_{31}x_1 + w_{32}x_2)}$$

Πού δείχνει το  $W_1 = (w_{11}, w_{12})$

Προς τα κόκκινα σημεία. Γιατί;

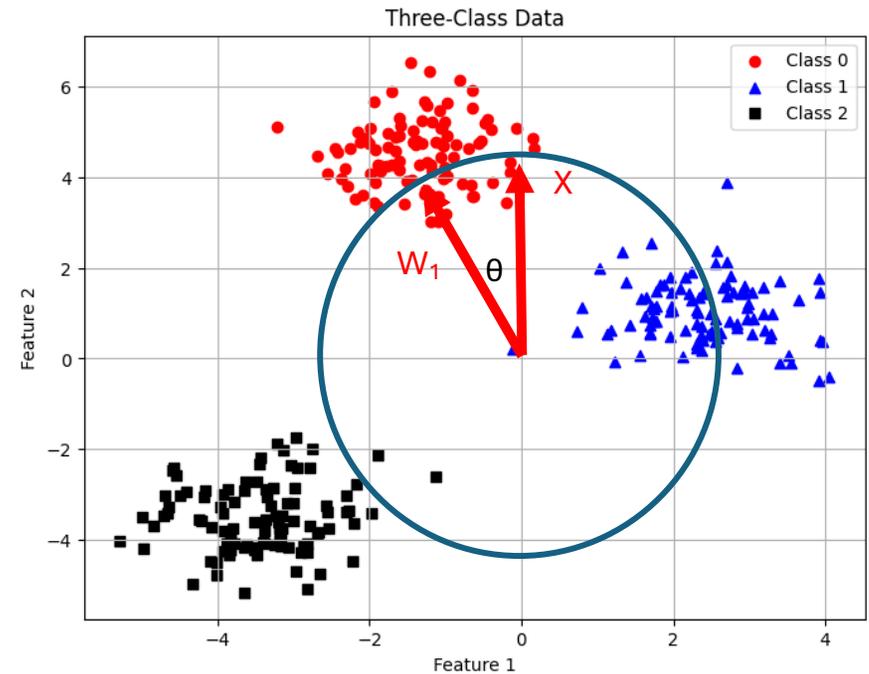


$$y_1 = \frac{\exp(w_{11}x_1 + w_{12}x_2)}{\exp(w_{11}x_1 + w_{12}x_2) + \exp(w_{21}x_1 + w_{22}x_2) + \exp(w_{31}x_1 + w_{32}x_2)}$$

Πού δείχνει το  $W_1 = (w_{11}, w_{12})$

Προς τα κόκκινα σημεία. Γιατί;

**Μικρή** γωνία  $\theta \rightarrow$  **μεγάλο** εσωτερικό  
γινόμενο  $\rightarrow$  **μεγάλος** αριθμητής  $\rightarrow$   
**μεγάλο**  $y_1$

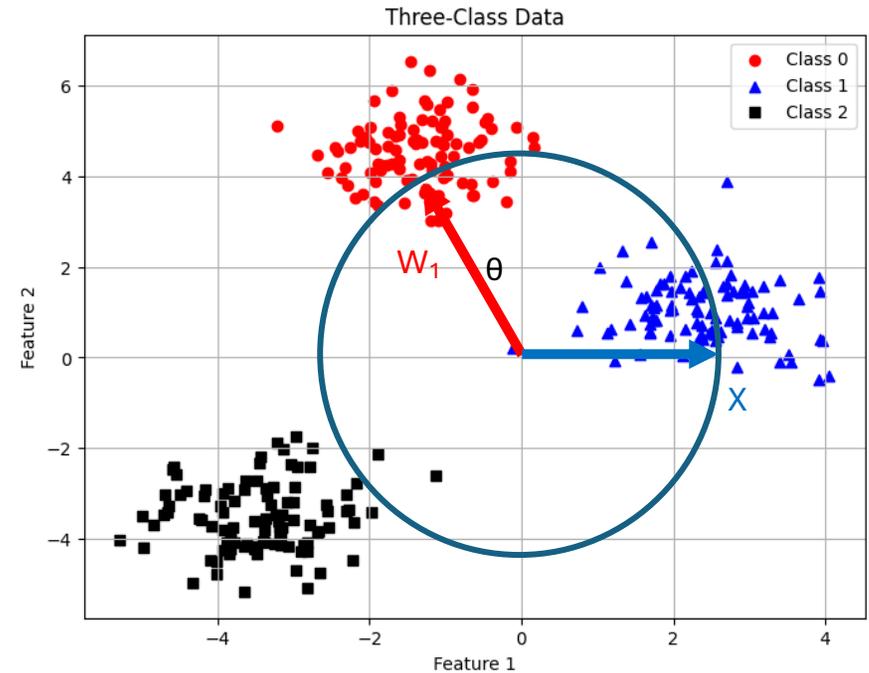


$$y_1 = \frac{\exp(w_{11}x_1 + w_{12}x_2)}{\exp(w_{11}x_1 + w_{12}x_2) + \exp(w_{21}x_1 + w_{22}x_2) + \exp(w_{31}x_1 + w_{32}x_2)}$$

Πού δείχνει το  $W_1 = (w_{11}, w_{12})$

Προς τα κόκκινα σημεία. Γιατί;

**Μεγάλη** γωνία  $\theta \rightarrow$  **μικρό** εσωτερικό  
γινόμενο  $\rightarrow$  **μικρός** αριθμητής  $\rightarrow$   
**μικρό**  $y_1$

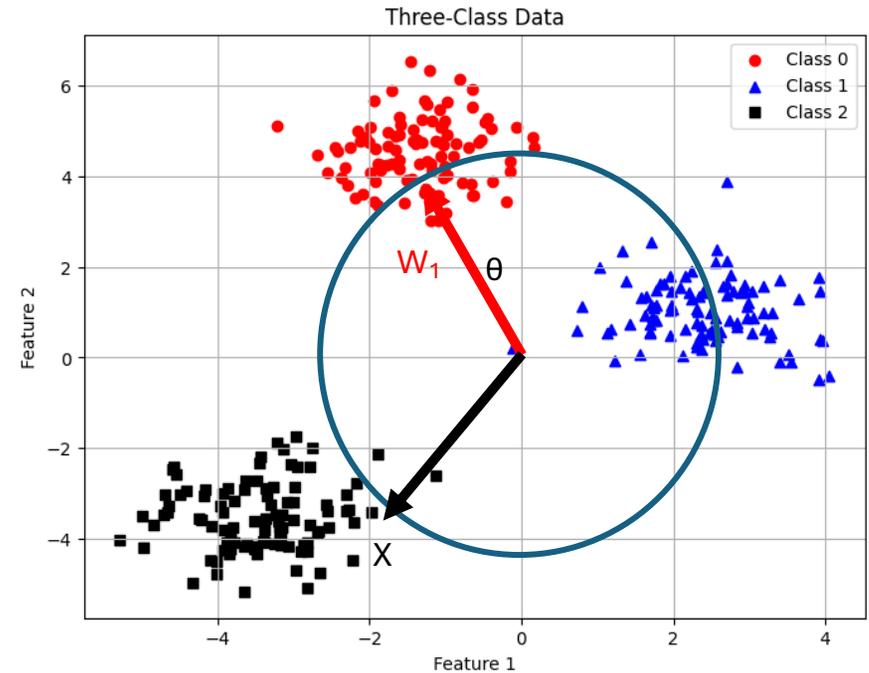


$$y_1 = \frac{\exp(w_{11}x_1 + w_{12}x_2)}{\exp(w_{11}x_1 + w_{12}x_2) + \exp(w_{21}x_1 + w_{22}x_2) + \exp(w_{31}x_1 + w_{32}x_2)}$$

Πού δείχνει το  $W_1 = (w_{11}, w_{12})$

Προς τα κόκκινα σημεία. Γιατί;

**Μεγάλη** γωνία  $\theta \rightarrow$  **μικρό** εσωτερικό  
γινόμενο  $\rightarrow$  **μικρός** αριθμητής  $\rightarrow$   
**μικρό**  $y_1$



$$y_1 = \frac{\exp(w_{11}x_1 + w_{12}x_2)}{\exp(w_{11}x_1 + w_{12}x_2) + \exp(w_{21}x_1 + w_{22}x_2) + \exp(w_{31}x_1 + w_{32}x_2)}$$

# Πολυδιάστατα Παραδείγματα - Διανύσματα

---

$$W = (1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0)$$

$$X_1 = (0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1)$$

$$X_2 = (1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0, 0)$$

$$X_3 = (0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0)$$

Ποια γωνία  $\theta$   
είναι μικρότερη

- $\angle WX_1$
- $\angle WX_2$
- $\angle WX_3$

# Πολυδιάστατα Παραδείγματα - Διανύσματα

---

$$W = (1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0)$$

$$X_1 = (0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1)$$

$$X_2 = (1, -1, 0, 0, 1, -1, 0, 0, 1, -1, 0, 0)$$

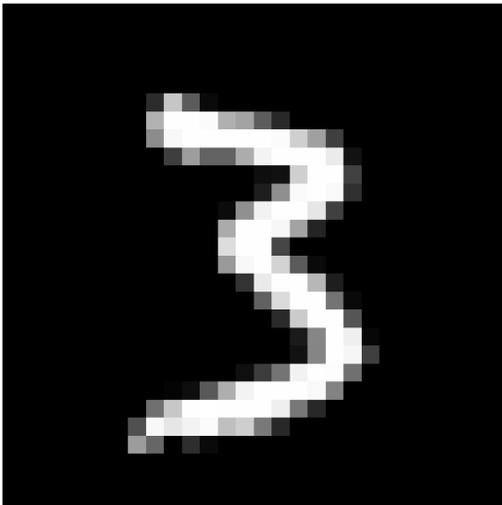
$$X_3 = (0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0)$$

Ποια γωνία  $\theta$   
είναι μικρότερη

- $\angle WX_1$
- $\angle WX_2$
- $\angle WX_3$

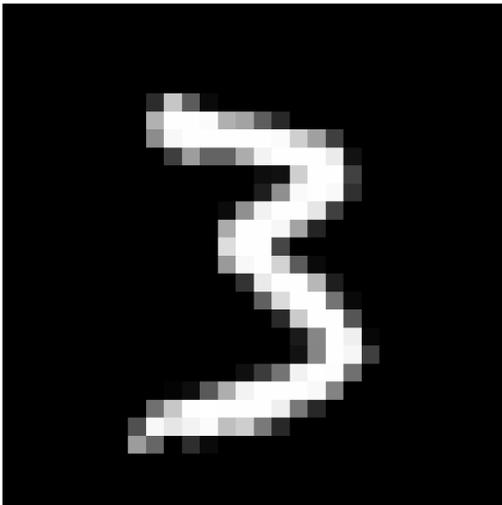
# Πολυδιάστατα Παραδείγματα - Εικόνες

---



# Πολυδιάστατα Παραδείγματα - Εικόνες

---



$28 \times 28 = 784$  πίξελ

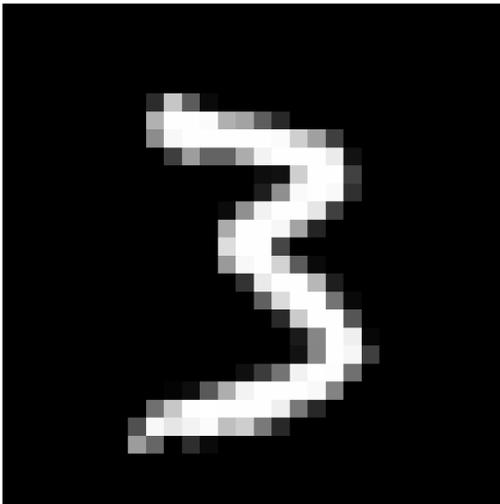
μαύρο πίξελ: 0

άσπρο πίξελ: 255

γκρίζο πίξελ: (0,255)

# Πολυδιάστατα Παραδείγματα - Εικόνες

---

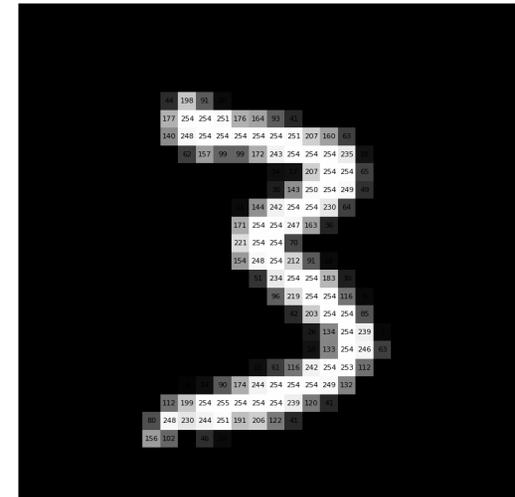


28 x 28 = 784 πίξελ

μαύρο πίξελ: 0

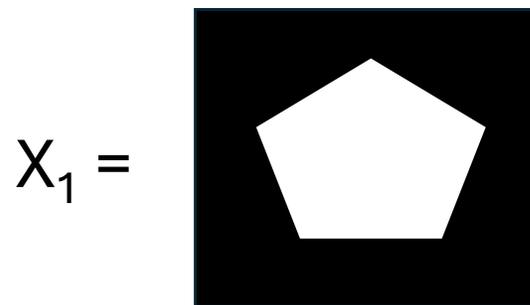
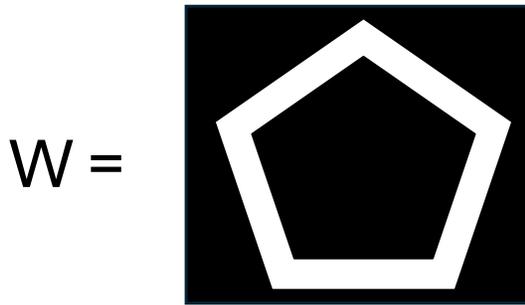
άσπρο πίξελ: 255

γκρίζο πίξελ: (0,255)



# Πολυδιάστατα Παραδείγματα - Εικόνες

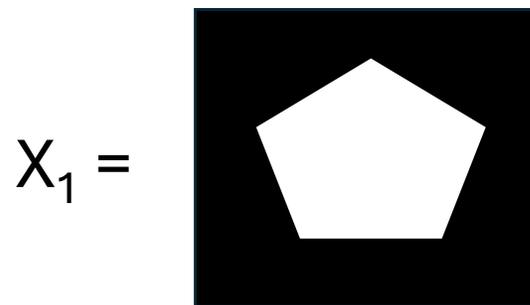
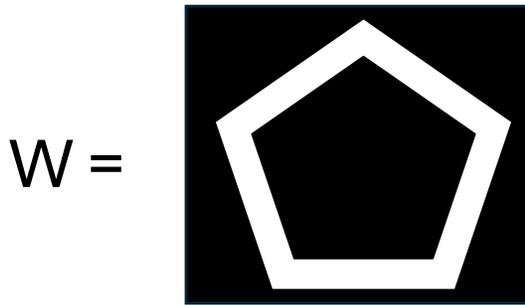
---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$

# Πολυδιάστατα Παραδείγματα - Εικόνες

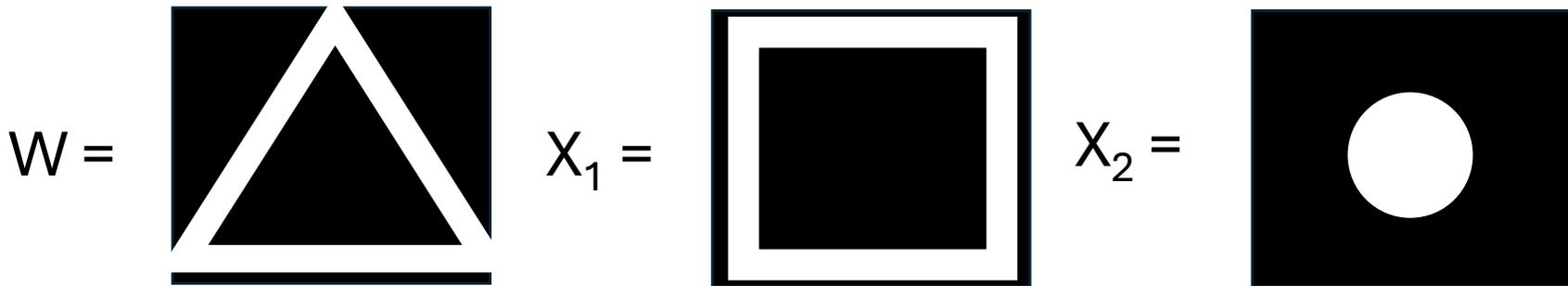
---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$

# Πολυδιάστατα Παραδείγματα - Εικόνες

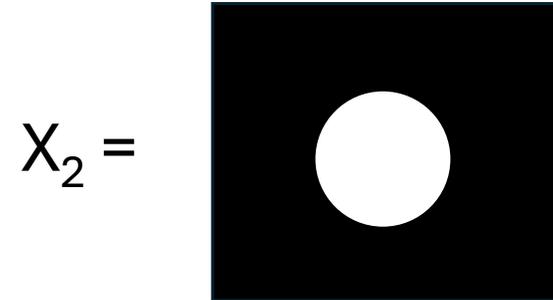
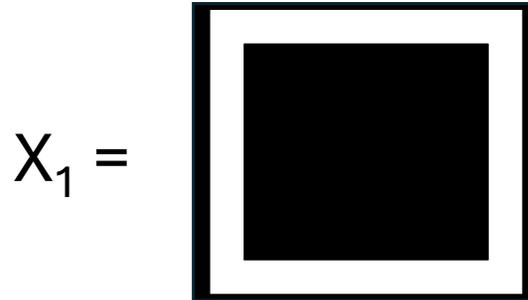
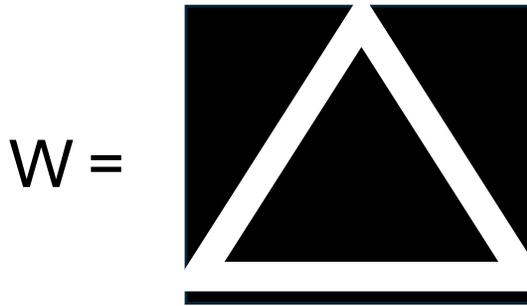
---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$

# Πολυδιάστατα Παραδείγματα - Εικόνες

---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$



Έγχρωμες εικόνες: **RGB** –  
**κόκκινα (R)**, **πράσινα (G)**  
και **μπλε (B)** πίξελ

Original



Red Channel



Green Channel

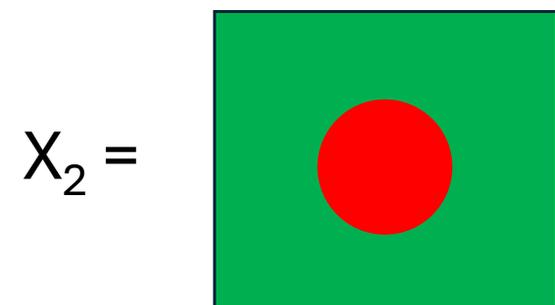
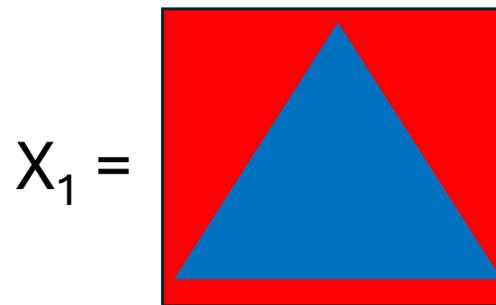
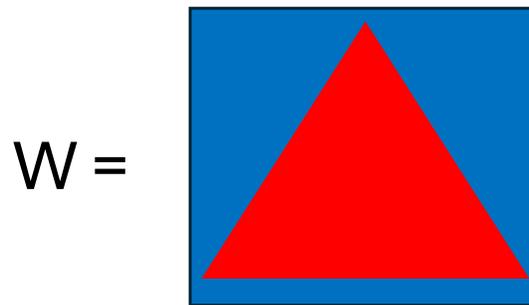


Blue Channel



# Πολυδιάστατα Παραδείγματα - Εικόνες

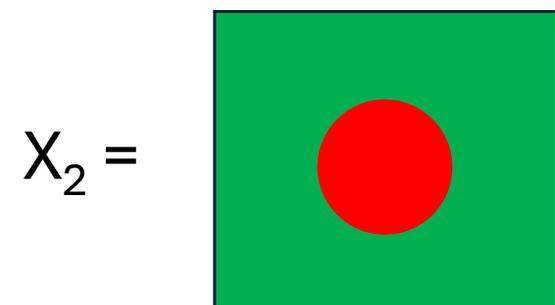
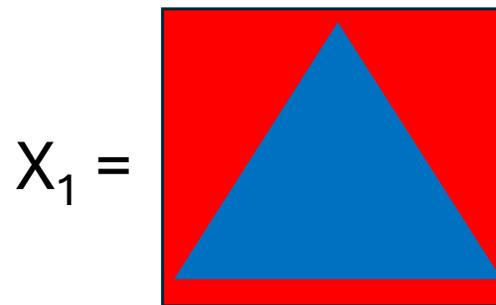
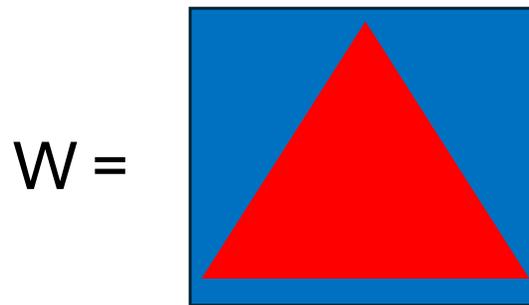
---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$

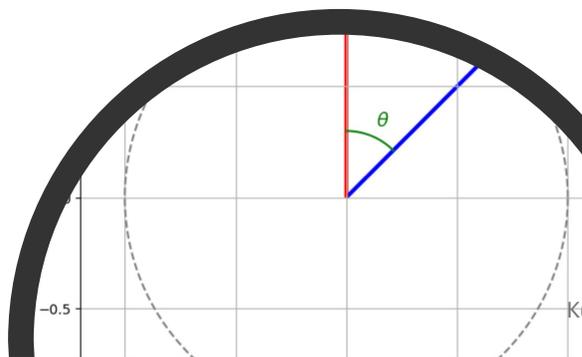
# Πολυδιάστατα Παραδείγματα - Εικόνες

---



Ποια γωνία  $\theta$  είναι μικρότερη:  $\angle WX_1$  ή  $\angle WX_2$

## Το Εσωτερικό Γινόμενο και η Γεωμετρία της Λογιστικής Παλινδρόμησης



Το Εσωτερικό Γινόμενο  $\langle W, X \rangle$  είναι μεγάλο όταν το  $W$  και το  $X$ :

- «Μοιάζουν»
- Είναι κοντά στον χώρο
- Σχηματίζουν μικρή γωνία  $\angle WX$

Το Εσωτερικό  
Γεωμετρία της  
Παλινδρόμησης

Το Εσωτερικό Γινόμενο  
 $\langle W, Y \rangle$  είναι μεγάλο όταν το  $W$

**Επόμενη Διάλεξη:** Συνδέουμε τη  
γεωμετρία του εσωτερικού γινομένου με  
τα νευρωνικά δίκτυα και τις  
ενσωματώσεις (embeddings)

