

Εισαγωγή στην Εκμάθηση Μηχανής

Μάθημα 2 - Supervised learning

Γιώργος Σφήκας



Σημαντικές έννοιες από το προηγούμενο μάθημα

- ❏ AI vs Machine Learning vs Deep Learning
- ❏ Supervised vs Unsupervised vs Reinforcement learning
- ❏ Μοντέλο vs Μέθοδος
- ❏ Classification vs Regression
- ❏ Discrete vs Continuous

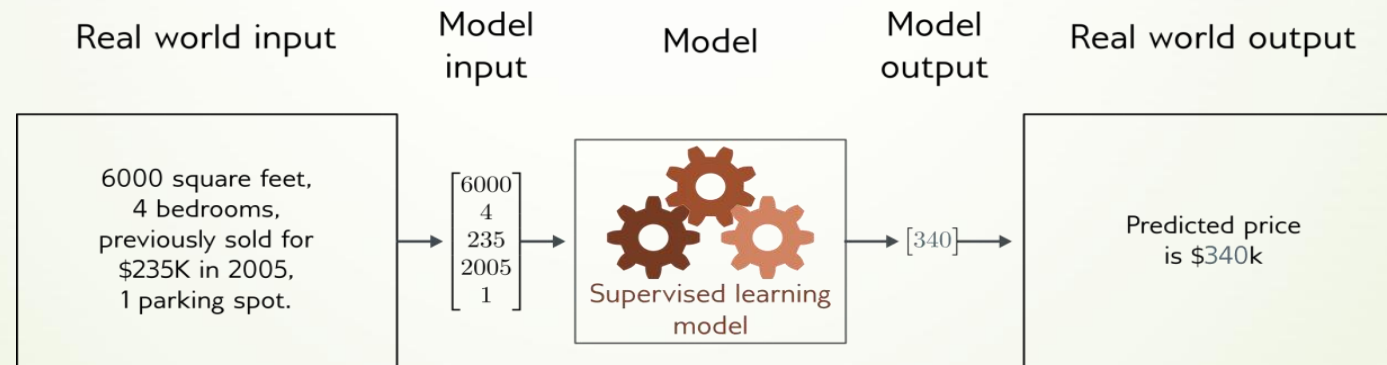


Τι θα δούμε σήμερα

- ✂ Επισκόπηση των συστατικών ενός συστήματος εκμάθησης με επίβλεψη - Supervised Learning
- ✂ Μια εφαρμογή με ένα απλό παράδειγμα

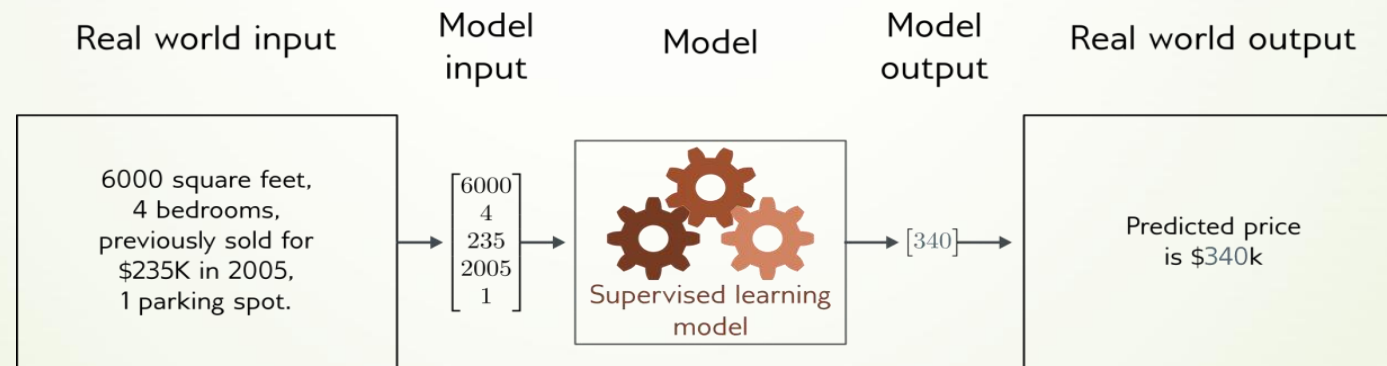
Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

- Supervised learning model = ορίζει **απεικόνιση** από μία ή περισσότερες **εισόδους** σε μία ή περισσότερες **εξόδους**



Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

- ❏ Supervised learning model = ορίζει **απεικόνιση** από μία ή περισσότερες **εισόδους** σε μία ή περισσότερες **εξόδους**
- ❏ Νέες έννοιες σε αυτό το μάθημα:
 - ❏ Συμπερασμός (Inference)
 - ❏ Παράμετροι ενός μοντέλου (parameters)
 - ❏ Εκπαίδευση ή εκμάθηση ενός μοντέλου (training / learning a model)





Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

$$y = f[x]$$

x, y είναι διανύσματα. π.χ. στο προηγούμενο παράδειγμα, $x \in \mathbb{R}^5$, $y \in \mathbb{R}$

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

Πιο πλήρης
διατύπωση

$$y = f[x, \varphi]$$

x, y και φ
είναι
διανύσματα

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

«*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$y = f[x, \varphi]$$

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ «*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$y = f[x, \varphi]$$

❏ “Βέλτιστες” = επιθυμούμε να παίρνουμε ένα λογικό αποτέλεσμα y για οποιαδήποτε είσοδο x

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ «*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$y = f[x, \varphi]$$

- ❏ “Βέλτιστες” = επιθυμούμε να παίρνουμε ένα λογικό αποτέλεσμα y για οποιαδήποτε είσοδο x
- ❏ Πως όμως θα ορίσουμε τι είναι “λογικό αποτέλεσμα y για οποιαδήποτε είσοδο x ” ;

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

 Απάντηση: Θα κατασκευάσουμε ένα μέτρο καταλληλότητας για παραμέτρους φ του μοντέλου μας



$$y = f[x, \varphi]$$

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

📌 Απάντηση: Θα κατασκευάσουμε ένα μέτρο καταλληλότητας για παραμέτρους φ του μοντέλου μας



$$y = f[x, \varphi]$$

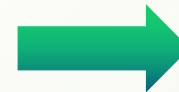
📌 ...χρησιμοποιώντας δύο συστατικά

📌 1. Ένα “σύνολο εκπαίδευσης” (training set):



$$\{x_i, y_i\}_{i=1..I}$$

📌 2. Μια “συνάρτηση κόστους” (cost function):



$$L[\varphi]$$

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ «*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]]$$

❏ Αφού βρούμε $\hat{\varphi}$, έχουμε ένα μοντέλο που δουλεύει καλά για το *training set*

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ «*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]]$$

- ❏ Αφού βρούμε $\hat{\varphi}$, έχουμε ένα μοντέλο που δουλεύει καλά για το *training set*
- ❏ Στην πράξη θέλουμε να δούμε πόσο καλά δουλεύει για εισόδους που δεν είναι μέρος του training set. Χρησιμοποιούμε ένα ξεχωριστό σύνολο που ονομάζουμε σύνολο ελέγχου (*test set*)

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ «*Learning*» or «*training*» a model σημαίνει την αναζήτηση βέλτιστων παραμέτρων φ

$$\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]]$$

- ❏ Αφού βρούμε $\hat{\varphi}$, έχουμε ένα μοντέλο που δουλεύει καλά για το *training set*
- ❏ Στην πράξη θέλουμε να δούμε πόσο καλά δουλεύει για εισόδους που δεν είναι μέρος του training set. Χρησιμοποιούμε ένα ξεχωριστό σύνολο που ονομάζουμε σύνολο ελέγχου (*test set*)
- ❏ Αν το μοντέλο δουλεύει καλά και για το test set (αυτό δεν είναι δεδομένο!), λέμε ότι έχει ικανότητα γενίκευσης (*generalization*)

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

▣ Παρένθεση -- ανακεφαλαίωση εννοιών

▣ Input και Outputs

▣ Model

▣ Model parameters

▣ Training και test set

▣ Cost ή Loss function

▣ Training ή learning (η διδασκαλία)

▣ Generalization

Εκμάθηση με επίβλεψη (supervised learning) - επισκόπηση

❏ Παρένθεση -- ανακεφαλαίωση εννοιών

- ❏ Input και Outputs
- ❏ Model
- ❏ Model parameters
- ❏ Training και test set
- ❏ Cost ή Loss function
- ❏ Training ή learning (η διδασκασία)
- ❏ Generalization

Μόλις περιγράψαμε τα
βασικά συστατικά και
“συνταγή” του machine
learning !

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

- ❑ Ας κάνουμε τα παραπάνω πιο χειροπιαστά με ένα απλό παράδειγμα!
- ❑ Μοντέλο που προβλέπει την τιμή μιας εξόδου ($y \in \mathbb{R}$) δοθείσης μιας εισόδου ($x \in \mathbb{R}$)

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

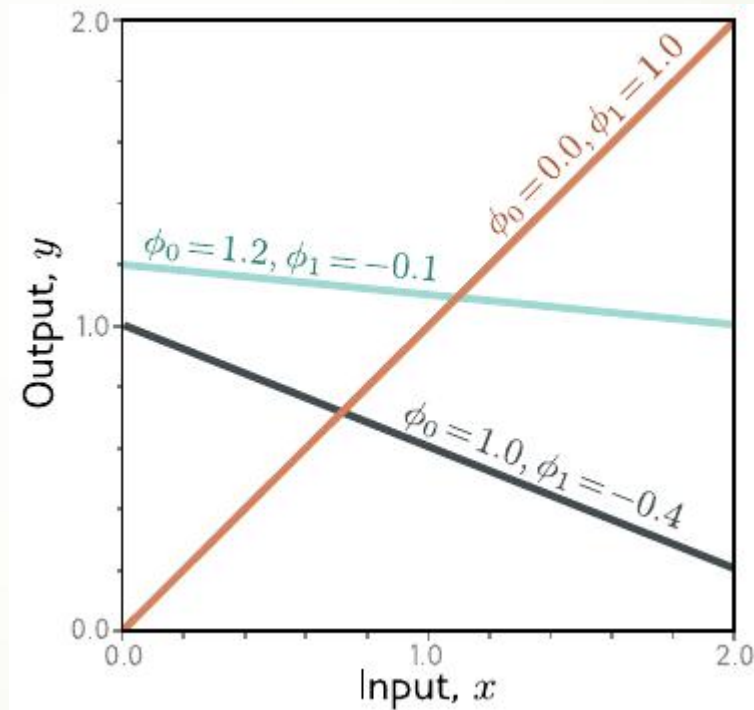
$$y = f[x, \varphi] = \varphi_0 + \varphi_1 x$$

$$\text{όπου } x \in \mathbb{R}, y \in \mathbb{R}, \varphi = \begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} \in \mathbb{R}^2$$

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

$$y = f[x, \varphi] = \varphi_0 + \varphi_1 x$$

$$\begin{aligned} \text{όπου } x &\in \mathbb{R}, y \in \mathbb{R}, \varphi = \\ &\begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} \in \mathbb{R}^2 \end{aligned}$$

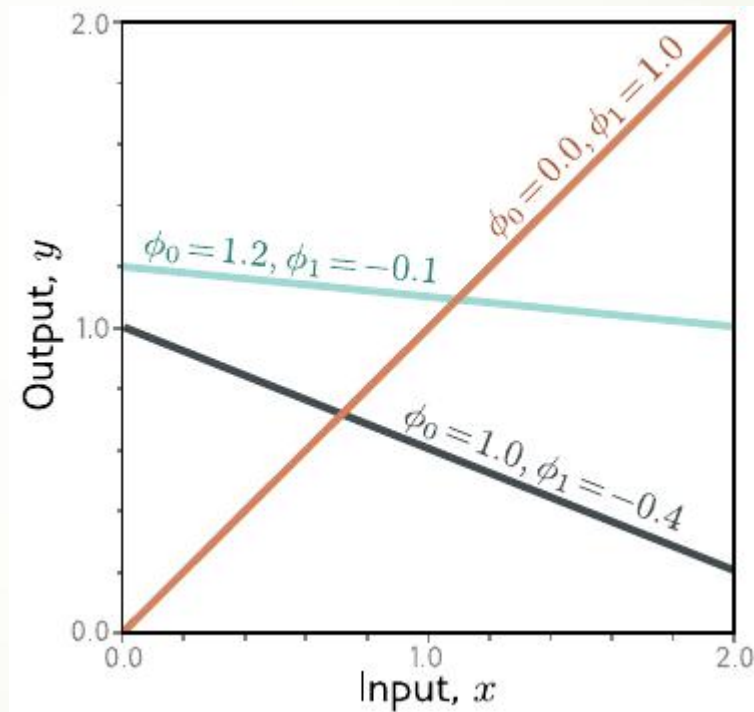


Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Αυτή η σχέση
ορίζει μια
οικογένεια
πιθανών
σχέσεων
εισόδου-εξόδου
(= όλες τις
πιθανές ευθείες)

$$y = f[x, \varphi] = \varphi_0 + \varphi_1 x$$

$$\begin{aligned} \text{όπου } x &\in \mathbb{R}, y \in \mathbb{R}, \varphi = \\ &\begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} \in \mathbb{R}^2 \end{aligned}$$

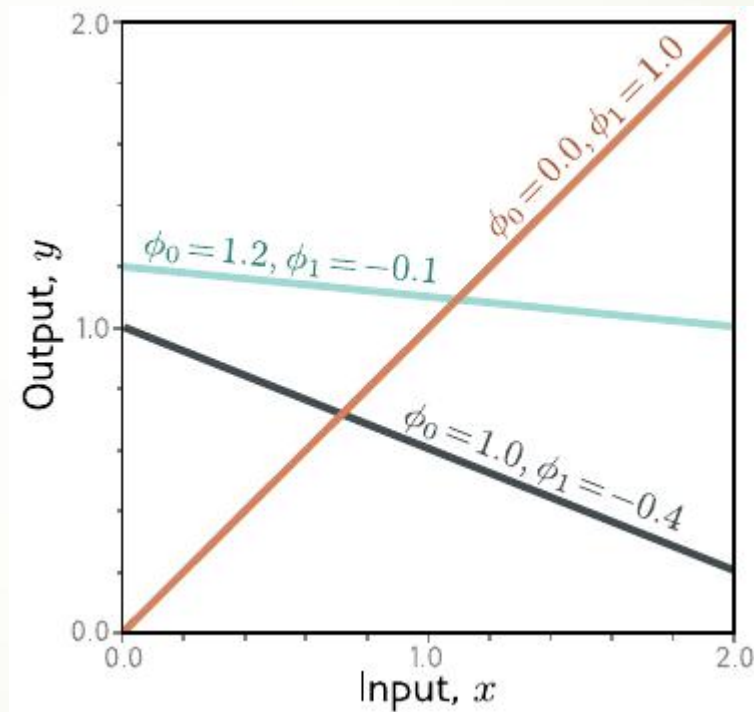


Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Αυτή η σχέση
ορίζει μια
οικογένεια
πιθανών
σχέσεων
εισόδου-εξόδου
(= όλες τις
πιθανές ευθείες)

$$y = f[x, \varphi] = \varphi_0 + \varphi_1 x$$

$$\begin{aligned} \text{όπου } x &\in \mathbb{R}, y \in \mathbb{R}, \varphi = \\ &\begin{bmatrix} \varphi_0 \\ \varphi_1 \end{bmatrix} \in \mathbb{R}^2 \end{aligned}$$



Η επιλογή
συγκεκριμένων
παραμέτρων
καθορίζει σε
ποιο μέλος της
οικογένειας
αναφερόμαστε
(= ποια ευθεία)

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Loss

❏ Loss (ή συνώνυμα cost):

❏ Ένα μέτρο της αποτυχίας του μοντέλου στις προβλέψεις του

$$\begin{aligned} \text{❏ } L[\varphi] &= \sum_{i=1}^I (f[x_i, \varphi] - y_i)^2 \\ &= \sum_{i=1}^I (\varphi_0 + \varphi_1 x_i - y_i)^2 \end{aligned}$$

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Loss

❏ Loss (ή συνώνυμα cost):

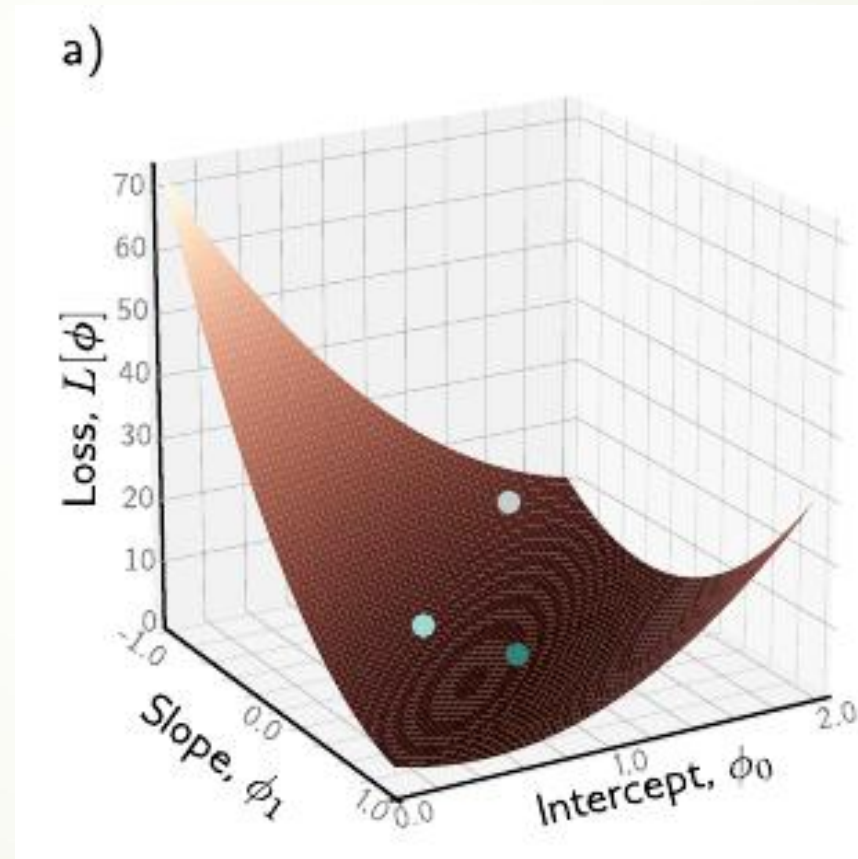
❏ Ένα μέτρο της αποτυχίας του μοντέλου στις προβλέψεις του

$$\begin{aligned} L[\varphi] &= \sum_{i=1}^I (f[x_i, \varphi] - y_i)^2 \\ &= \sum_{i=1}^I (\varphi_0 + \varphi_1 x_i - y_i)^2 \end{aligned}$$

«Least-squares»
loss

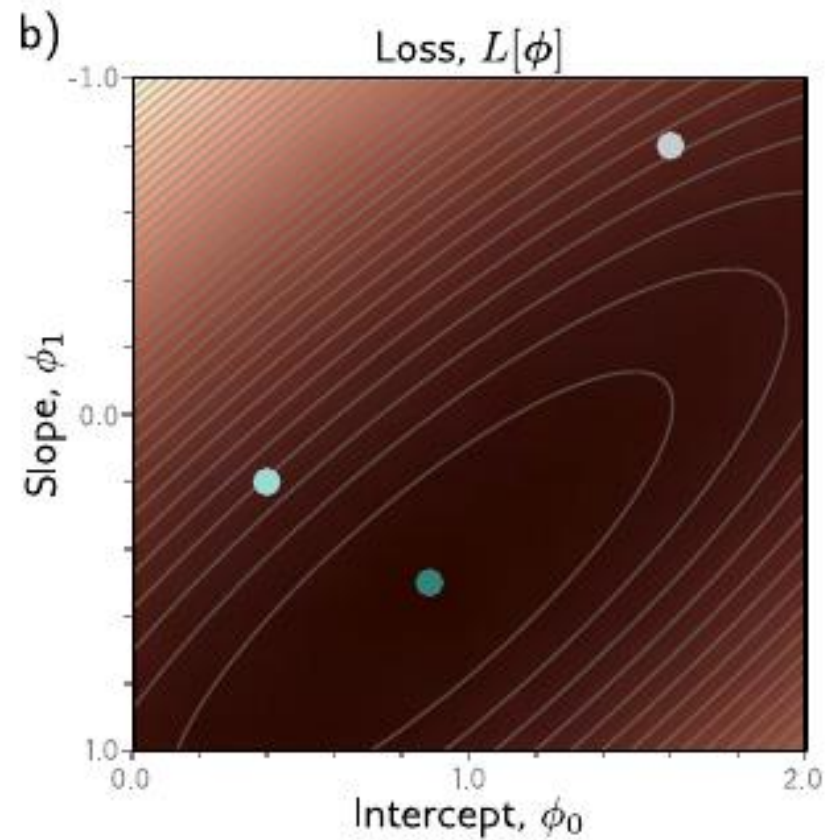
Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Loss



Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Loss



Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Training

- ❏ Η διαδικασία εύρεσης των παραμέτρων που ελαχιστοποιούν το loss ονομάζεται **model fitting, training** ή **learning** (el: ταίριασμα μοντέλου, εκπαίδευση, εκμάθηση).

π.χ. στο παράδειγμα: $\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]] = \underset{\varphi}{\operatorname{argmin}} \left[\sum_{i=1}^I (\varphi_0 + \varphi_1 x_i - y_i)^2 \right]$

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Training

- ❑ Η διαδικασία εύρεσης των παραμέτρων που ελαχιστοποιούν το loss ονομάζεται **model fitting, training** ή **learning** (el: ταίριασμα μοντέλου, εκπαίδευση, εκμάθηση).

π.χ. στο παράδειγμα: $\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]] = \underset{\varphi}{\operatorname{argmin}} \left[\sum_{i=1}^I (\varphi_0 + \varphi_1 x_i - y_i)^2 \right]$

- ❑ Η βασική λογική είναι:
 - ❑ 1. Επιλέγουμε παραμέτρους *τυχαία*

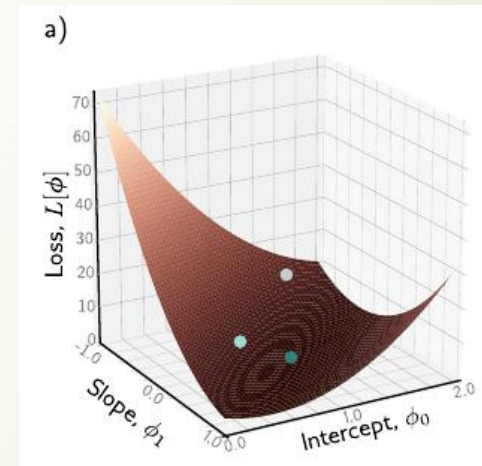
Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Training

- ❑ Η διαδικασία εύρεσης των παραμέτρων που ελαχιστοποιούν το loss ονομάζεται **model fitting, training** ή **learning** (el: ταίριασμα μοντέλου, εκπαίδευση, εκμάθηση).

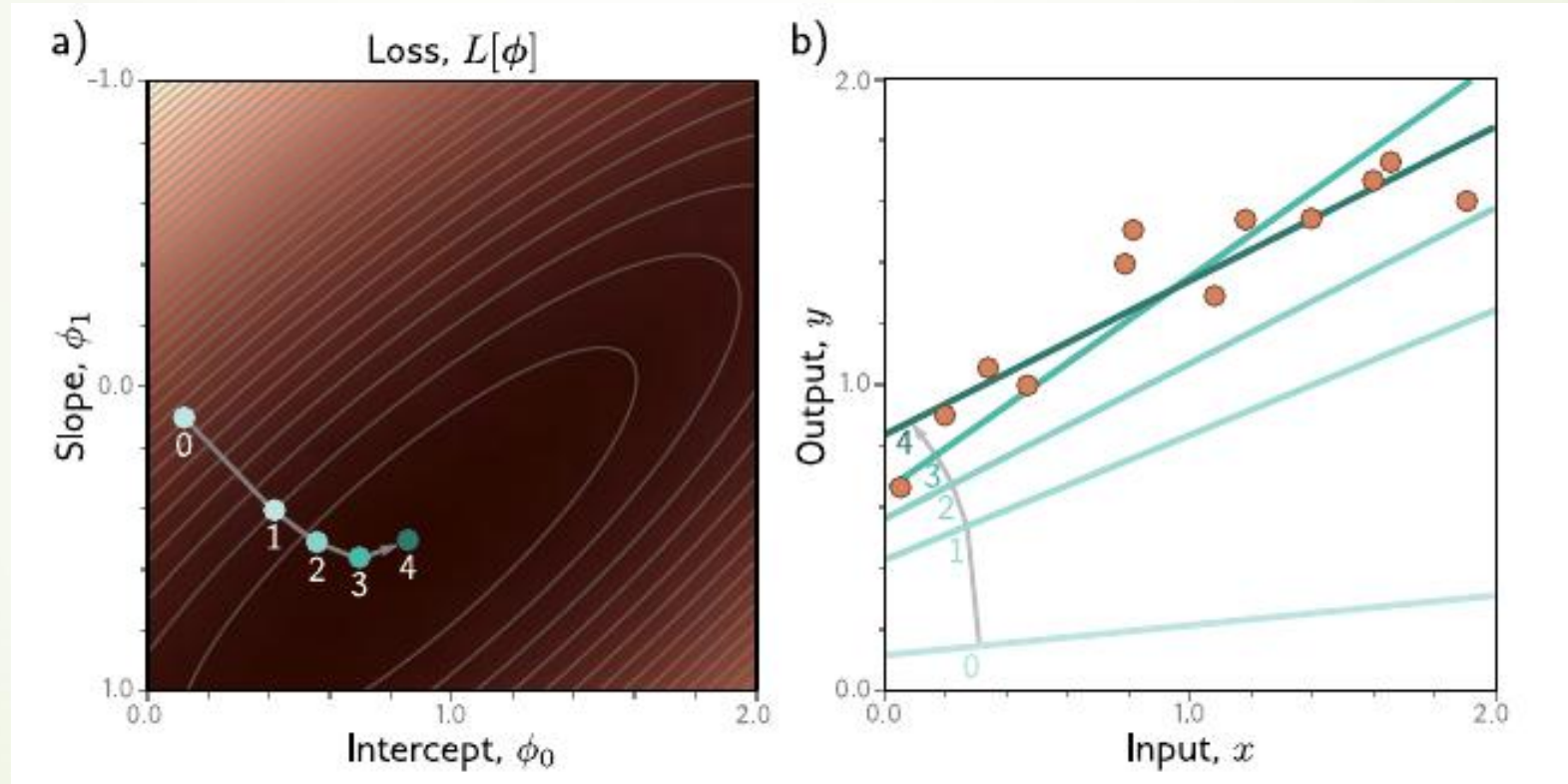
π.χ. στο παράδειγμα: $\hat{\varphi} = \underset{\varphi}{\operatorname{argmin}} [L[\varphi]] = \underset{\varphi}{\operatorname{argmin}} \left[\sum_{i=1}^I (\varphi_0 + \varphi_1 x_i - y_i)^2 \right]$

- ❑ Η βασική λογική είναι:
 - ❑ 1. Επιλέγουμε παραμέτρους *τυχαία*
 - ❑ 2. Βελτιώνουμε τις παραμέτρους
“περπατώντας” την επιφάνεια του loss προς “τα κάτω”
 - ❑ 3. Επιστροφή στο (2) μέχρι να μην μπορεί να υπάρξει βελτίωση



Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Training





Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Testing

- ❑ Με την διαδικασία του **testing** (εl: έλεγχος μοντέλου) ελέγχω κατά πόσο το μοντέλο μου δουλεύει καλά στον “πραγματικό κόσμο”
- ❑ Για το testing απλά υπολογίζω το loss για ένα δεύτερο, ξεχωριστό σύνολο δεδομένων (test set, ελληνικά: σύνολο ελέγχου)

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Testing

- ❑ Με την διαδικασία του **testing** (ει: έλεγχος μοντέλου) ελέγχω κατά πόσο το μοντέλο μου δουλεύει καλά στον “πραγματικό κόσμο”
- ❑ Για το testing απλά υπολογίζω το loss για ένα δεύτερο, ξεχωριστό σύνολο δεδομένων (test set, ελληνικά: σύνολο ελέγχου)
- ❑ Ο βαθμός που το μοντέλο μου κάνει καλές ή κακές προβλέψεις εξαρτάται από...

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Testing

- ❑ Με την διαδικασία του **testing** (εl: έλεγχος μοντέλου) ελέγχω κατά πόσο το μοντέλο μου δουλεύει καλά στον “πραγματικό κόσμο”
- ❑ Για το testing απλά υπολογίζω το loss για ένα δεύτερο, ξεχωριστό σύνολο δεδομένων (test set, ελληνικά: σύνολο ελέγχου)
- ❑ Ο βαθμός που το μοντέλο μου κάνει καλές ή κακές προβλέψεις εξαρτάται από:
 - ❑ 1. Την επιλογή του μοντέλου - κατά πόσο είναι κατάλληλο για το πρόβλημα; Κατά πόσο εκφραστικό - πλούσιο είναι το μοντέλο (πόσο ευρεία είναι η οικογένεια αντιστοιχίσεων που αναπαριστά);
 - ❑ 2. Πόσο καλά έχω εκπαιδεύσει το μοντέλο = Έχω βρει το ελάχιστο της L ;
 - ❑ 3. Πόσο αντιπροσωπευτικό και πλήρες είναι το training set που έχω στη διάθεσή μου;

Παράδειγμα γραμμικής παλινδρόμησης (linear regression)

Testing

- ❏ Θέλουμε να αποφύγουμε το **underfitting**, όπως και το **overfitting**
- ❏ Όταν το μοντέλο μας δουλεύει καλά και για το test set, λέμε ότι έχει καλή ικανότητα γενίκευσης (**generalization**) ή ότι γενικεύει τις προβλέψεις του καλά στο test set

Ανακεφαλαίωση

- ❏ Ένα supervised learning model είναι μια συνάρτηση $y = f[x, \varphi]$, η οποία απεικονίζει τις εισόδους x στις εξόδους y .
- ❏ Οι παράμετροι φ καθορίζουν τον ακριβή τρόπο της απεικόνισης
- ❏ Για να εκπαιδεύσουμε το μοντέλο, βρίσκουμε παραμέτρους που ελαχιστοποιούν την συνάρτηση κόστους $L[\varphi]$ για ένα σύνολο ζευγαριών $\{x_i, y_i\}_{i=1..I}$.
- ❏ Η συνάρτηση κόστους ποσοτικοποιεί πόσο απέχει η κάθε πρόβλεψη $f[x_i, \varphi]$ του μοντέλου από την πραγματικότητα y_i
- ❏ Υπολογίζουμε το κόστος και σε ένα ξεχωριστό σύνολο, για να αξιολογήσουμε την ικανότητα γενίκευσης σε “νέα” δεδομένα / δεδομένα που δεν έχει “δει”.

..και τι θα δούμε στα επόμενα μαθήματα

- ❏ Το μοντέλο που είδαμε είναι γραμμικό - θα δούμε ένα (πολύ απλό) μη-γραμμικό μοντέλο (Κεφάλαιο 3)
- ❏ Από τα “Ρηχά” στα “Βαθιά” νευρωνικά δίκτυα (Κεφάλαιο 4)
- ❏ Θα μιλήσουμε για άλλους τρόπους να περιγράψουμε το «Loss» (Κεφάλαιο 5)
- ❏ Θα εξετάσουμε καλύτερα τον τρόπο εκπαίδευσης (Κεφάλαια 6,7)
- ❏ Θα εξετάσουμε καλύτερα τον τρόπο αποτίμησης της επίδοσης ενός μοντέλου (Κεφάλαιο 8)
- ❏ Θα δούμε τεχνικές “κανονικοποίησης” (Κεφάλαιο 9)
- ❏ Η “αιχμή” της έρευνας: Σύγχρονα Συνελικτικά δίκτυα, Υπολειμματικά δίκτυα, Μετασχηματιστές (Κεφάλαια 10,11,12)
- ❏ Τα βασικά της μη-επιβλεπόμενης μάθησης (Κεφάλαιο 14)
- ❏ Παραγωγικά ανταγωνιστικά δίκτυα (Κεφάλαιο 15)



Στο εργαστήριο..

- 📺 Βασικές γνώσεις στη θεωρία, με απλές εφαρμογές σε Python

Ερωτήσεις ... ;;;

