MEDLAB
MEDICAL TECHNOLOGY & INTELLIGENT INFORMATION SYSTEMS

Τεχνητή Νοημοσύνη και Επιστήμη των Υλικών



- Διδάσκων: Δημήτριος Ι. Φωτιάδης
- Τμήμα Μηχανικών Επιστήμης Υλικών
- Website: https://medlab.cc.uoi.gr/

- Ιωάννινα 2026 - 2027

Περιεχόμενα



1. Εισαγωγή στα Γραμμικά Μοντέλα
2. Απλή Γραμμική Παλινδρόμηση
3. Πολλαπλή Γραμμική Παλινδρόμηση
4. Κανονικοποίηση: Ridge και Lasso
5. Λογιστική Παλινδρόμηση και Ταξινόμηση
6. Μη Γραμμικά Μοντέλα μέσω Συναρτήσεων Πυρήνα
7. Εφαρμογές στην Επιστήμη Υλικών και Σύνοψη

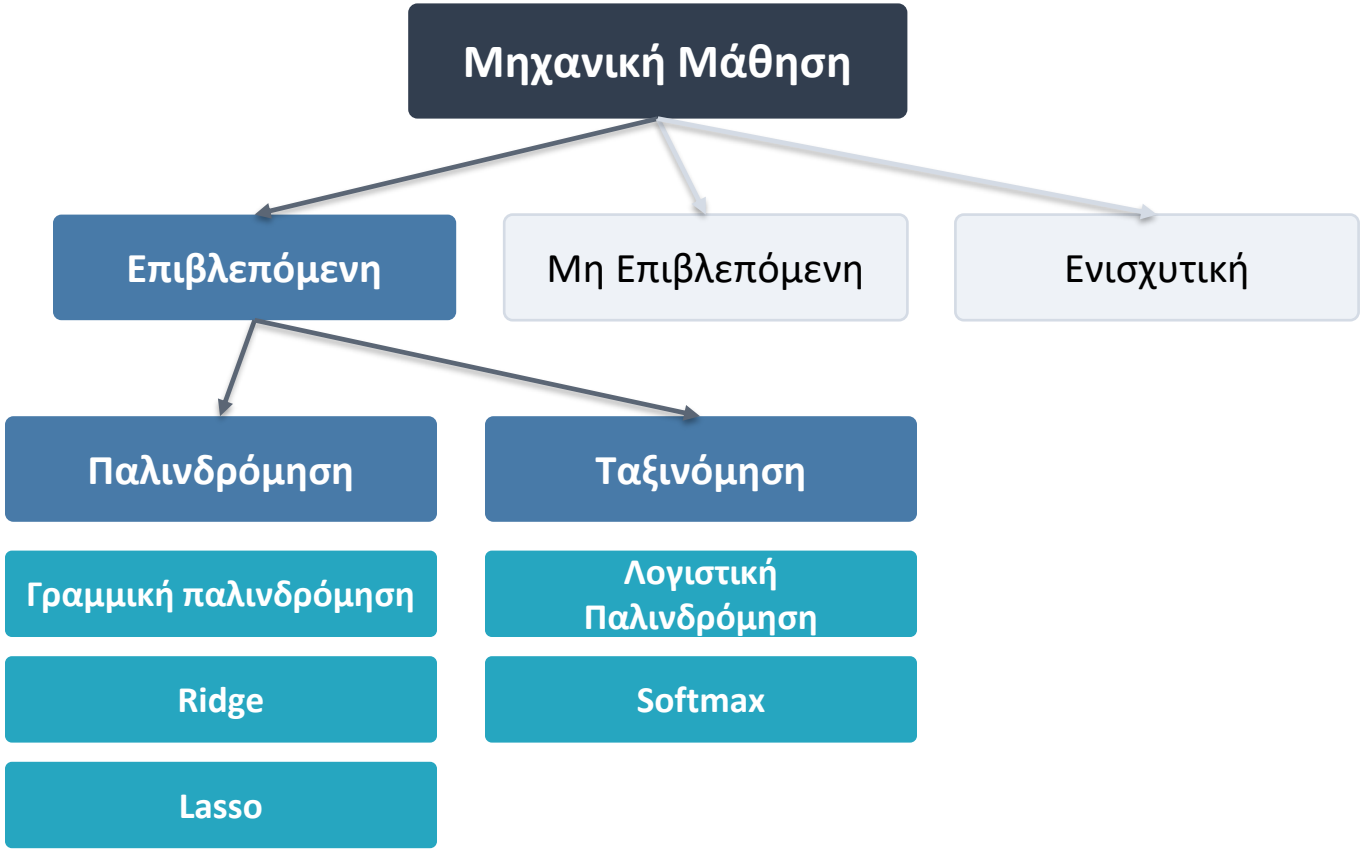
Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.



Ενότητα 1

Εισαγωγή στα Γραμμικά Μοντέλα

Ανακεφαλαίωση: τα γραμμικά μοντέλα



Τι θυμόμαστε από τη Διάλεξη 1

- **Δεδομένα:** πίνακας X (δείγματα × χαρακτηριστικά) και ετικέτες y
- **Μοντέλο:** μια μαθηματική συνάρτηση y = f(x)
- **Εκπαίδευση:** εύρεση των παραμέτρων που ελαχιστοποιούν τη συνάρτηση απώλειας
- **Γενίκευση:** καλή απόδοση σε νέα δεδομένα, όχι μόνο στα δεδομένα εκπαίδευσης
- **Υπερπροσαρμογή:** το μοντέλο «μαθαίνει» και τον θόρυβο. Σήμερα θα δούμε μια λύση: την κανονικοποίηση

Στην έρευνα Kaggle 2020, η γραμμική και η λογιστική παλινδρόμηση ήταν οι πιο διαδεδομένοι αλγόριθμοι Μηχανικής Μάθησης

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Αντικείμενο και μαθησιακοί στόχοι

Αντικείμενο της Διάλεξης

Τα γραμμικά μοντέλα: η γραμμική παλινδρόμηση, η κανονικοποίηση (Ridge, Lasso), η λογιστική παλινδρόμηση για ταξινόμηση και οι συναρτήσεις πυρήνα για μη γραμμικά προβλήματα.

AI for Eng.

Βασικό βοήθημα

Z. Liu, Artificial Intelligence for Engineers, Κεφ. 3: Linear Models

Στο τέλος της διάλεξης θα μπορείτε να:

1. Διατυπώνετε ένα πρόβλημα ως γραμμική παλινδρόμηση
2. Εξηγείτε τη μέθοδο ελαχίστων τετραγώνων και την αναλυτική λύση
3. Χρησιμοποιείτε Ridge και Lasso κατά της υπερπροσαρμογής
4. Εφαρμόζετε λογιστική παλινδρόμηση σε προβλήματα ταξινόμησης
5. Εξηγείτε πώς οι συναρτήσεις πυρήνα κάνουν ένα γραμμικό μοντέλο μη γραμμικό

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

NumPy, pandas, SciPy: τι κάνει η καθεμία

Τρεις **βιβλιοθήκες** της Python, δηλαδή έτοιμα «πακέτα» κώδικα.
Τις χρησιμοποιούμε σε κάθε ανάλυση δεδομένων.

NumPy
Numerical Python

Πίνακες αριθμών και γρήγορες πράξεις πάνω τους (άθροισμα, μέσος όρος, πολλαπλασιασμός πινάκων).

pandas
Python Data Analysis

Πίνακες δεδομένων με **ονόματα στηλών**, σαν φύλλο Excel. Διαβάζει αρχεία (.csv, .xlsx) και οργανώνει τα δεδομένα.

SciPy
Scientific Python

Έτοιμοι επιστημονικοί αλγόριθμοι: **βελτιστοποίηση**, στατιστική, προσαρμογή καμπυλών.

Και οι τρεις είναι δωρεάν και ανοιχτού κώδικα. Έρχονται έτοιμες στο Google Colab και στο Anaconda.

Πώς συνδέονται

scikit-learn, matplotlib : έτοιμα μοντέλα Μηχανικής Μάθησης και γραφήματα

pandas: οργανώνει τα δεδομένα

SciPy: λύνει τα προβλήματα

NumPy: κρατά τους αριθμούς, το κοινό θεμέλιο όλων

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πώς μοιάζουν στην πράξη

NumPy

```
import numpy as np
hv = np.array([180, 235, 310])
hv.mean()      # 241.7
```

180	235	310

Ένας πίνακας αριθμών (1D).
Ο μέσος όρος σε μία εντολή.

pandas

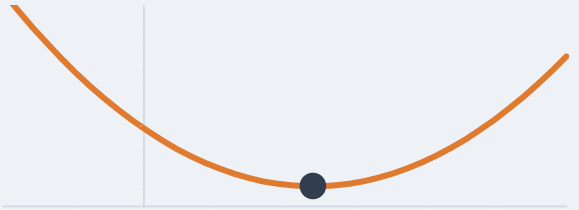
```
import pandas as pd
df = pd.read_csv("alloys.csv")
df.head()      # οι πρώτες γραμμές
```

	%C	%Cr	T	HV
0	0,20	1,0	800	180
1	0,40	1,5	850	235
2	0,60	1,0	900	310

Πίνακας με ονόματα στηλών (DataFrame).

SciPy

```
from scipy.optimize import minimize
res = minimize(loss, w0)
res.x          # τα βάρη
```



Βρίσκει το **ελάχιστο** της απώλειας = τα βάρη.

Πότε χρησιμοποιούμε την κάθε βιβλιοθήκη;

- Αριθμοί και πράξεις → **NumPy**
- Δεδομένα με στήλες → **pandas**
- Βελτιστοποίηση, στατιστική → **SciPy**

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Γιατί ξεκινάμε από τα γραμμικά μοντέλα

Τα γραμμικά μοντέλα είναι από τους **πιο απλούς**, αλλά **πολύ χρήσιμους**, αλγορίθμους Μηχανικής Μάθησης

Απλά

Λίγες παράμετροι (w, b). Για τη βασική μορφή υπάρχει **αναλυτική λύση**, χωρίς επαναληπτική βελτιστοποίηση.

Ερμηνεύσιμα

Κάθε συντελεστής w δείχνει **πόσο αλλάζει η έξοδος** όταν αλλάζει ένα χαρακτηριστικό, π.χ. η σκληρότητα ανά 0,1% C.

Λίγα δεδομένα

Χρειάζονται **λιγότερα δείγματα** από τα σύνθετα μοντέλα. Σημαντικό στα υλικά, όπου κάθε δείγμα κοστίζει.

Βάση για άλλους

Ο γραμμικός συνδυασμός **w·x + b** ξαναεμφανίζεται στις SVM και σε κάθε νευρώνα ενός νευρωνικού δικτύου.

Η πορεία της διάλεξης

Γραμμική παλινδρόμηση →	**Ridge & Lasso** →	**Λογιστική παλινδρόμηση** →	**Συναρτήσεις πυρήνα**
Πρόβλεψη συνεχών τιμών	*Λιγότερη υπερπροσαρμογή*	*Ταξινόμηση*	*Μη γραμμικά προβλήματα*

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Αλγόριθμος και μοντέλο

Ο όρος «γραμμικό μοντέλο» αναφέρεται στην πραγματικότητα στον **αλγόριθμο** που μας δίνει το μοντέλο. Τον κρατάμε επειδή έτσι χρησιμοποιείται στη βιβλιογραφία της Μηχανικής Μάθησης

Αλγόριθμος
y = w·x + b
w, b: άγνωστα

+

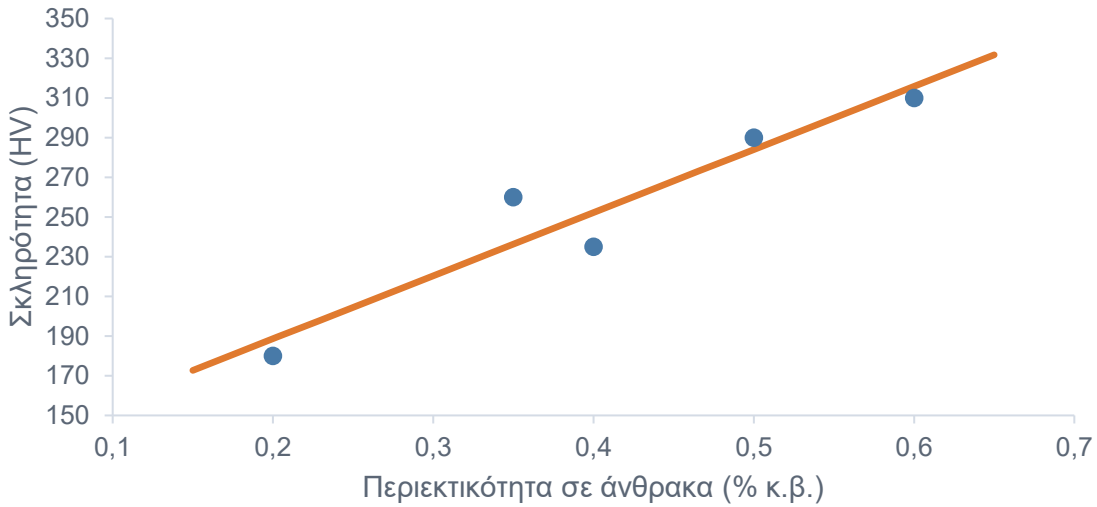
Δεδομένα
5 κράματα:
%C και σκληρότητα (HV)

→ ***Εκπαίδευση***

Μοντέλο
HV = 318·(%C) + 125
w, b: σταθερές τιμές

→

Πρόβλεψη
νέο κράμα με 0,45% C
→ ≈ 268 HV



Αλγόριθμος ≈ η απόφαση «θα προσαρμόσω ένα πολυώνυμο στα δεδομένα μου»

Μοντέλο ≈ το πολυώνυμο με **συγκεκριμένους συντελεστές**, όπως προέκυψαν από την προσαρμογή

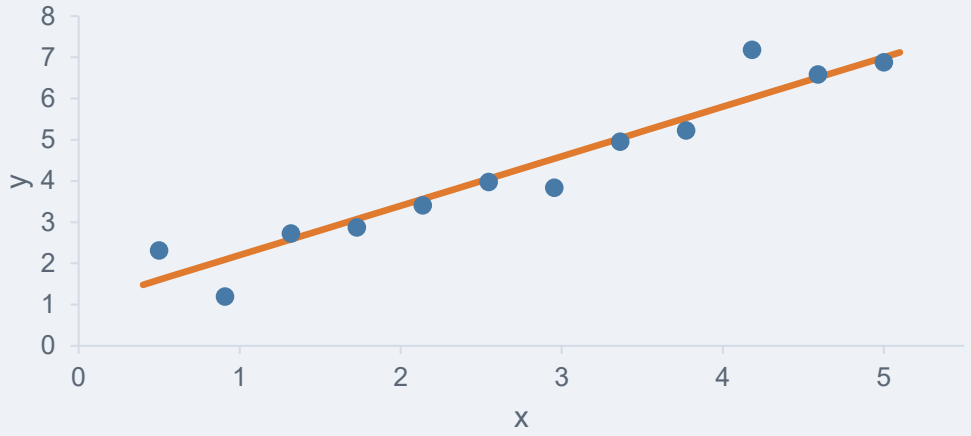
Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Γραμμικό μοντέλο: στενή και ευρεία έννοια

Στενή έννοια

Η έξοδος y είναι **γραμμικός συνδυασμός** των χαρακτηριστικών x_1, x_2, … της εισόδου.

$$y = w_1x_1 + w_2x_2 + \ldots + w_Jx_J + b$$

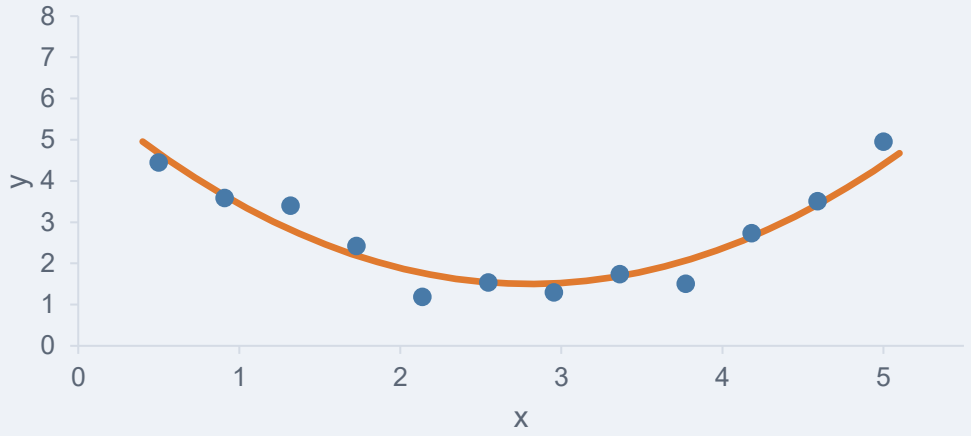


Π.χ. σκληρότητα ως γραμμική συνάρτηση της σύστασης.

Ευρεία έννοια

Με **συναρτήσεις πυρήνα** (kernel functions) το γραμμικό μοντέλο αποκτά **μη γραμμικές** δυνατότητες.

$$y = w_1x + w_2x^2 + b$$



Γραμμικό μοντέλο στις νέες μεταβλητές x και x^2.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Τρεις διαφορετικοί όροι

⚠ Οι τρεις όροι σε κάποιες περιπτώσεις σημαίνουν **το ίδιο**. Η σύγχυση οφείλεται στο ότι τους χρησιμοποιούν **διαφορετικοί κλάδοι**, π.χ. πληροφορική, μαθηματικά, μηχανική.

Γραμμικό μοντέλο

Linear model

Συνήθης όρος στην Μηχανική Μάθηση. Οικογένεια αλγορίθμων για **παλινδρόμηση** και **ταξινόμηση**.

Γραμμική παλινδρόμηση

Linear regression

Συνήθης όρος στην στατιστική. Πρόβλεψη μιας **συνεχούς** μεταβλητής από άλλες μεταβλητές (predictors → response).

Προσαρμογή καμπύλης

Curve fitting

Συνήθης όρος στη μηχανική και τα πειράματα. Εύρεση της καμπύλης που περνά **πιο κοντά** από τις μετρήσεις.

Σε αυτή την διάλεξη:

- Τα γραμμικά μοντέλα από τη σκοπιά της Μηχανικής Μάθησης: **παλινδρόμηση** και **ταξινόμηση**
- Απλά μοντέλα, άρα **αναλυτική λύση** σε κλειστή μορφή

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.



Ενότητα 2

Απλή Γραμμική Παλινδρόμηση

Τι είναι η παλινδρόμηση;

Η **παλινδρόμηση** (regression) βρίσκει τη σχέση ανάμεσα σε μια **συνεχή** μεταβλητή-στόχο και μία ή περισσότερες άλλες μεταβλητές, ώστε να μπορούμε να την **προβλέψουμε** ή να την **εξηγήσουμε**.

Οι δύο ρόλοι των μεταβλητών

Ανεξάρτητη (x)
αυτό που γνωρίζουμε ή ελέγχουμε

→

Εξαρτημένη (y)
αυτό που θέλουμε να βρούμε (συνεχής τιμή)

Π.χ. %C (x) → σκληρότητα σε HV (y)

Δύο στόχοι

Πρόβλεψη

Πόση θα είναι η σκληρότητα ενός **νέου** κράματος, πριν το φτιάξουμε;

Εξήγηση

Πόσο επηρεάζει ο άνθρακας τη σκληρότητα; Αυξάνει ή μειώνει;

Ερωτήματα παλινδρόμησης στα υλικά

- Σκληρότητα χάλυβα από την περιεκτικότητα σε άνθρακα
- Ρυθμός αποδόμησης κράματος Mg από την περιεκτικότητα σε Zn
- Μέτρο ελαστικότητας ικριώματος από το πορώδες
- Ποσοστό φαρμάκου που αποδεσμεύεται σε σχέση με τον χρόνο

*Κοινό σε όλα: η απάντηση είναι ένας **αριθμός** (πόσο;).*

Όταν η απάντηση είναι κατηγορία (ποιο;), μιλάμε για ταξινόμηση.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πρώτο βήμα: το διάγραμμα διασποράς

Πριν από κάθε μοντέλο, **σχεδιάζουμε τα δεδομένα**: κάθε δείγμα είναι ένα σημείο (x, y). Το διάγραμμα διασποράς (scatter plot) δείχνει αν υπάρχει σχέση και τι μορφή έχει.

Θετική γραμμική

x ↑ → y ↑



π.χ. %C → σκληρότητα

Αρνητική γραμμική

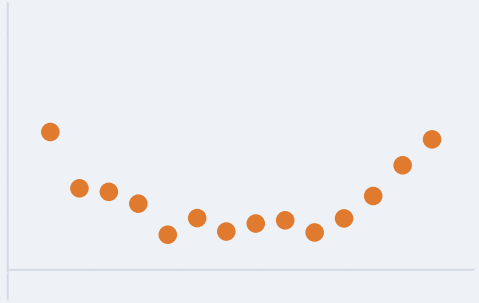
x ↑ → y ↓



π.χ. πορώδες → μέτρο ελαστικότητας

Μη γραμμική

όχι ευθεία



η ευθεία δεν αρκεί

Καμία σχέση

τυχαία διασπορά



το x δεν βοηθά να προβλέψουμε το y

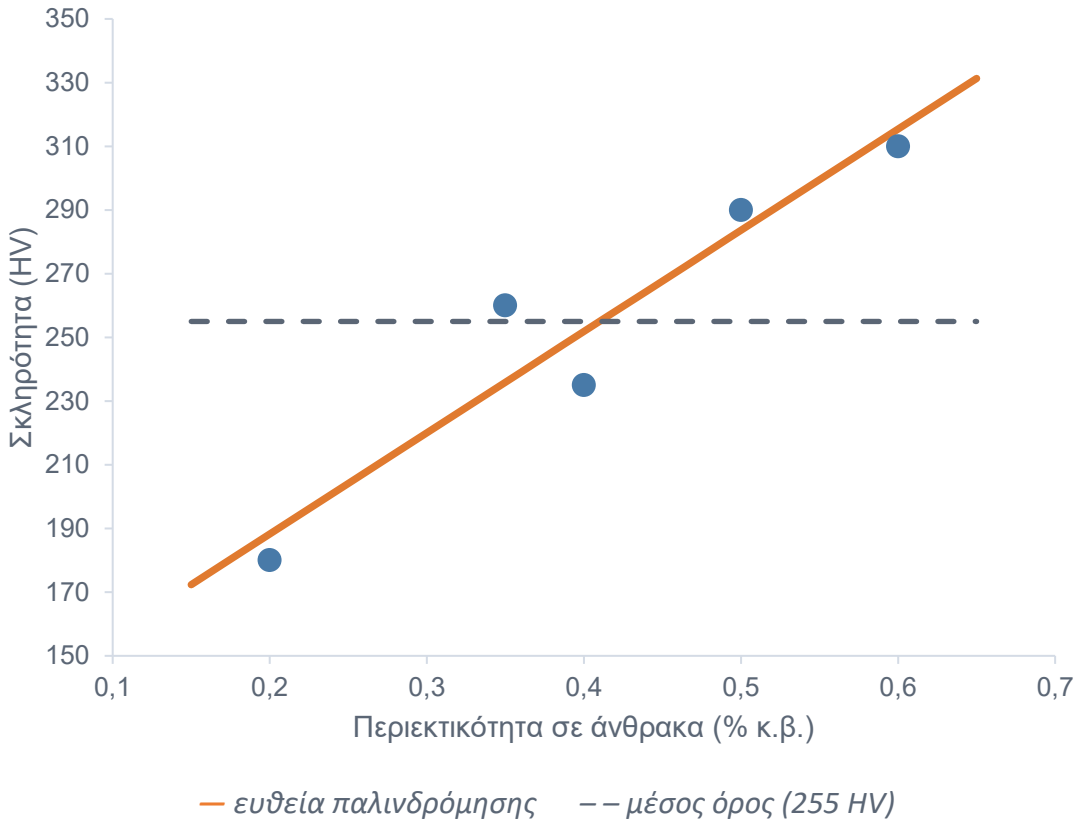
Τι μας λέει το διάγραμμα

Υπάρχει σχέση; | **Θετική ή αρνητική;** | **Γραμμική ή όχι;** | **Πόσο ισχυρή;** | **Ακραίες τιμές;**

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Τι σημαίνει η πρόβλεψη του μοντέλου

Η ευθεία **δεν περνά από όλα τα σημεία**. Δύο κράματα με το ίδιο %C μπορεί να έχουν διαφορετική σκληρότητα. Η πρόβλεψη είναι η **αναμενόμενη (μέση) τιμή**, όχι μια βέβαιη τιμή.



Το «μοντέλο βάσης»

Χωρίς καμία μεταβλητή, η καλύτερη πρόβλεψη είναι ο **μέσος όρος** (255 HV). Η ευθεία αξίζει μόνο αν προβλέπει **καλύτερα** από αυτόν.

$$y = w \cdot x + b + \varepsilon$$

ε: ό,τι δεν εξηγεί το μοντέλο

Ο όρος ε (σφάλμα ή θόρυβος) οφείλεται σε:

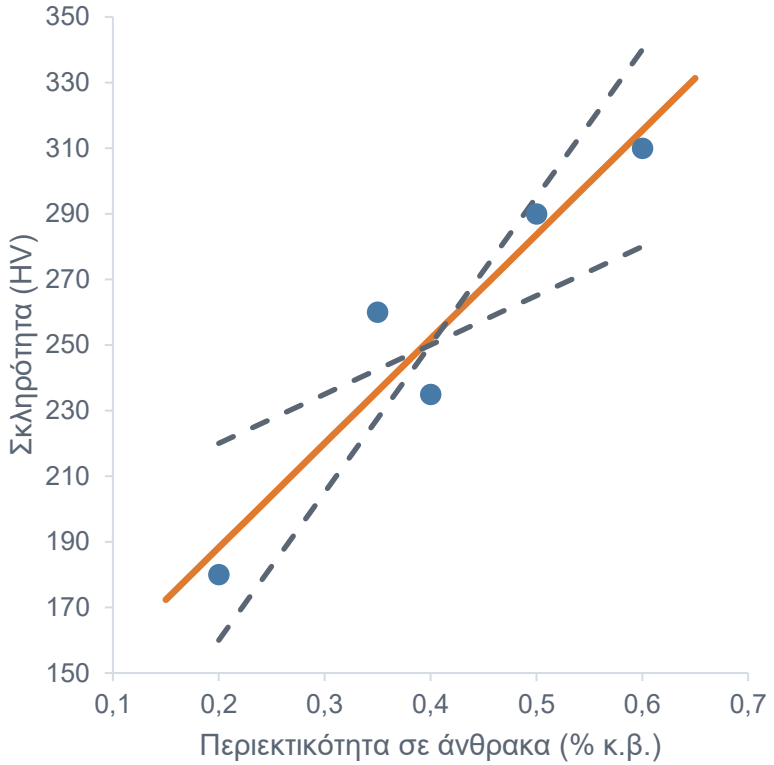
- **μεταβλητές που λείπουν** (π.χ. Cr, θερμική κατεργασία) → πολλαπλή παλινδρόμηση (Ενότητα 3)
- **μη γραμμική σχέση** → Ενότητα 6
- σφάλματα μέτρησης

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Το πρόβλημα με μία μεταβλητή

Ξεκινάμε από την πιο απλή περίπτωση: **μία είσοδος x** και **μία έξοδος y**. Έχουμε μετρήσεις (x_i, y_i), $i = 1, \ldots, l$, και ψάχνουμε την ευθεία που τις περιγράφει **καλύτερα** (Liu, §3.2.1).

Κράμα	%C (x)	HV (y)
M1	0,20	180
M2	0,35	260
M3	0,40	235
M4	0,50	290
M5	0,60	310



Ποια από τις ευθείες είναι η «καλύτερη»;

Τρία ερωτήματα

1. **Ποια μορφή έχει το μοντέλο;**
   → *ευθεία $f(x) = w \cdot x + b$*
2. **Πώς μετράμε το «καλύτερα»;**
   → *υπόλοιπα και τετράγωνα*
3. **Πώς βρίσκουμε τα w και b;**
   → *ελάχιστα τετράγωνα*

Δεδομένα: το παράδειγμα σκληρότητας της Διάλεξης 1 (ενδεικτικές τιμές).

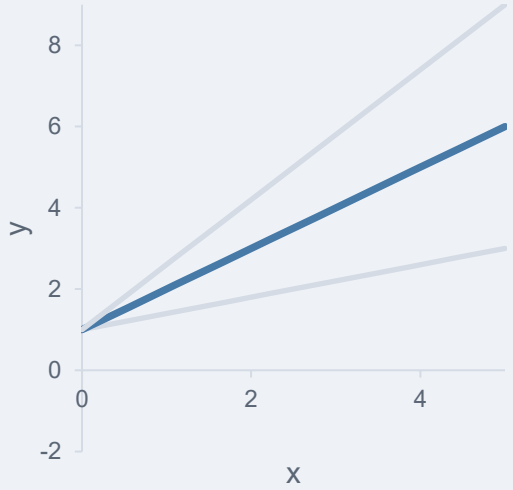
Η συνάρτηση f(x) = w·x + b

Το απλό γραμμικό μοντέλο έχει **δύο παραμέτρους**. Η εκπαίδευση σημαίνει να βρούμε τις τιμές τους από τα δεδομένα.

$$f(x) = w \cdot x + b$$

w: κλίση (slope)

Πόσο αλλάζει το y όταν το x αυξάνεται κατά 1.

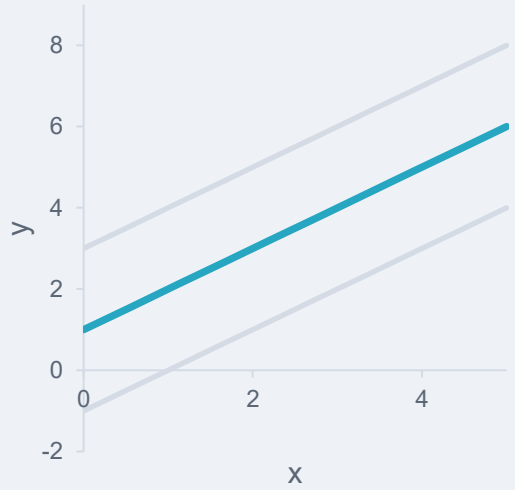


μεγαλύτερο w → πιο απότομη

Στο παράδειγμα: w ≈ 318 HV ανά 1% C, δηλαδή ≈ **+32 HV ανά 0,1% C**.

b: σταθερός όρος (intercept)

Η τιμή του y όταν x = 0. Μετακινεί την ευθεία πάνω ή κάτω.



μεγαλύτερο b → ψηλότερα

Στο παράδειγμα: b ≈ 125 HV. Προσοχή: το x = 0 είναι **έξω από το εύρος των δεδομένων**.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Ανεξάρτητες και εξαρτημένες μεταβλητές

Τα ίδια μεγέθη έχουν **διαφορετικά ονόματα** ανάλογα με το πεδίο.
Αξίζει να τα αναγνωρίζετε όλα, γιατί θα τα συναντήσετε στη βιβλιογραφία.

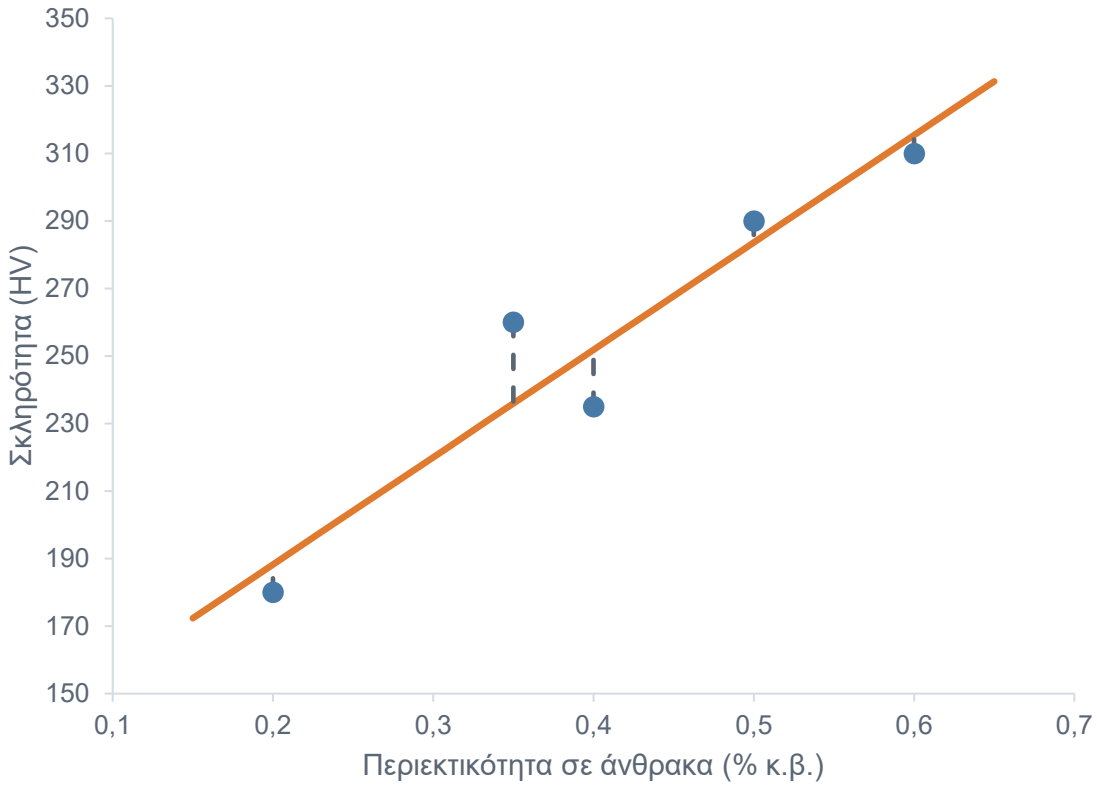
	x: είσοδος	**y: έξοδος**
Μηχανική Μάθηση	χαρακτηριστικό (attribute / feature)	ετικέτα (label)
Μαθηματικά	ανεξάρτητη μεταβλητή	εξαρτημένη μεταβλητή
Στατιστική	προβλεπτική μεταβλητή (predictor)	μεταβλητή απόκρισης (response)
Παράδειγμα υλικών	%C, θερμοκρασία, χρόνος κατεργασίας	σκληρότητα (HV)

*Προσοχή: «ανεξάρτητη» μεταβλητή **δεν** σημαίνει ότι τα χαρακτηριστικά είναι ανεξάρτητα μεταξύ τους.*

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Υπόλοιπα (residuals)

Πόσο απέχει το μοντέλο από τις μετρήσεις;



Οι διακεκομμένες γραμμές είναι τα **υπόλοιπα**: η κατακόρυφη απόσταση κάθε μέτρησης από την ευθεία.

- **Πρόβλεψη:** $f(x_i) = w \cdot x_i + b$
- **Υπόλοιπο:** $e_i = y_i - f(x_i)$
- $e_i > 0$: η μέτρηση είναι **πάνω** από την ευθεία
- $e_i < 0$: η μέτρηση είναι **κάτω** από την ευθεία

%C	y (HV)	f(x) (HV)	e (HV)
0,20	180	188,2	-8,2
0,35	260	235,9	+24,1
0,40	235	251,8	-16,8
0,50	290	283,6	+6,4
0,60	310	315,4	-5,4

Συνήθως μετράμε την απόσταση με την **Ευκλείδεια** απόσταση, δηλαδή με τα **τετράγωνα** των υπολοίπων.

Μέθοδος ελαχίστων τετραγώνων (OLS)

Η «καλύτερη» ευθεία είναι αυτή με το **μικρότερο άθροισμα τετραγώνων** των υπολοίπων.

Η μέθοδος λέγεται ελάχιστα τετράγωνα (ordinary least squares, OLS).

$$(w^*, b^*) = \arg\min \Sigma (w \cdot x_i + b - y_i)^2$$

«Βρες τα w και b που κάνουν το άθροισμα των τετραγώνων ελάχιστο»

Τύποι για μία μεταβλητή

$$w = \Sigma(x_i - \bar{x})(y_i - \bar{y}) / \Sigma(x_i - \bar{x})^2$$

$$b = \bar{y} - w \cdot \bar{x}$$

Γιατί τετράγωνα;

Δεν αλληλοαναιρούνται

Θετικά και αρνητικά υπόλοιπα δεν «σβήνουν» το ένα το άλλο.

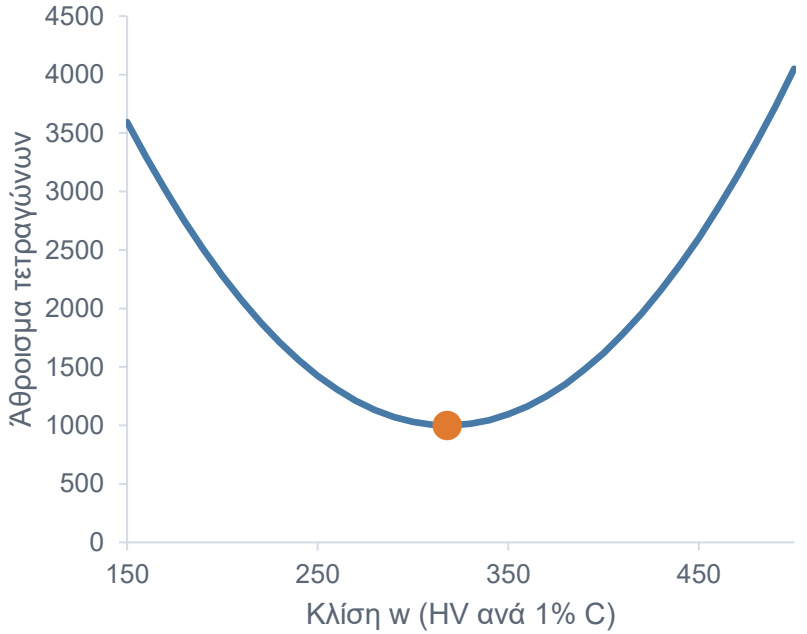
Τιμωρούν τα μεγάλα σφάλματα

Ένα υπόλοιπο 2 μετράει 4, ένα υπόλοιπο 10 μετράει 100.

Δίνουν αναλυτική λύση

Η συνάρτηση είναι ομαλή και έχει ένα ελάχιστο, που βρίσκεται με παραγώγους.

Μειονέκτημα: λόγω των τετραγώνων, η OLS είναι **ευαίσθητη σε έκτροπες τιμές** (outliers).



*Για το παράδειγμα σκληρότητας: το άθροισμα των τετραγώνων έχει **ένα ελάχιστο**, στο w ≈ 318.*



Ενότητα 3

Πολλαπλή Γραμμική Παλινδρόμηση

Γιατί δεν αρκεί μία μεταβλητή;

Μια ιδιότητα σπάνια εξαρτάται από **έναν μόνο** παράγοντα. Η σκληρότητα ενός χάλυβα εξαρτάται από τον άνθρακα, αλλά και από τα κραματικά στοιχεία και τη θερμική κατεργασία.

Πολλαπλή γραμμική παλινδρόμηση: μία συνεχής εξαρτημένη μεταβλητή και **δύο ή περισσότερες** ανεξάρτητες.

Απλή

$$HV = w \cdot \%C + b$$

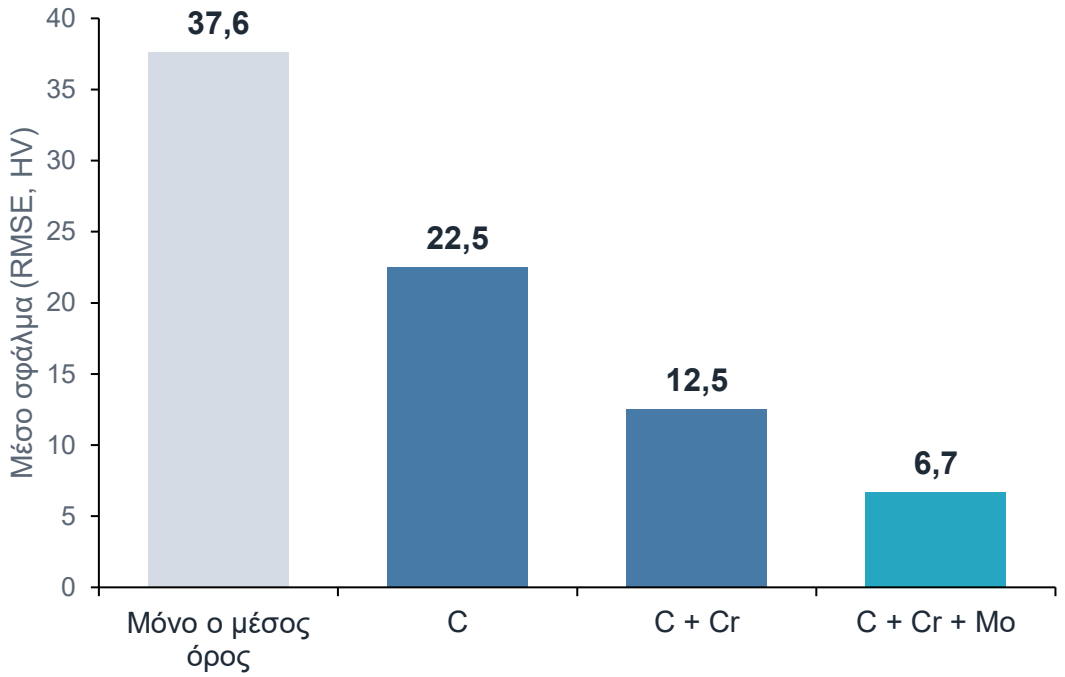
Ό,τι οφείλεται στο Cr ή στη θερμική κατεργασία «πέφτει» στο σφάλμα ε.

Πολλαπλή

$$HV = w_1 \cdot \%C + w_2 \cdot \%Cr + w_3 \cdot T + b$$

Κάθε παράγοντας παίρνει το δικό του βάρος. Οι συνεισφορές **προστίθενται**.

Όσο προσθέτουμε σχετικές μεταβλητές, το σφάλμα μειώνεται



Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πώς ερμηνεύουμε τα βάρη

Κάθε βάρος w_j δείχνει πόσο αλλάζει **κατά μέσο όρο** η έξοδος, όταν το χαρακτηριστικό j αυξηθεί κατά **1 μονάδα** και **όλα τα άλλα μείνουν σταθερά**.

$$HV = 300\cdot(\%C) + 40\cdot(\%Cr) + 25\cdot(\%Mo) + 120$$

$w_1 = 300$	+0,1% C (με Cr, Mo σταθερά) → σκληρότητα **+30 HV** κατά μέσο όρο
$w_2 = 40$	+1% Cr (με C, Mo σταθερά) → **+40 HV**
$w_3 = 25$	+1% Mo (με C, Cr σταθερά) → **+25 HV**
$b = 120$	τιμή όταν όλα είναι 0: εκτός εύρους δεδομένων, χωρίς φυσικό νόημα εδώ

Ενδεικτικό μοντέλο για επίδειξη, όχι πραγματικά δεδομένα.

Πρόσημο

+ : το χαρακτηριστικό αυξάνει την ιδιότητα

− : τη μειώνει

Προσοχή στη σύγκριση

Τα βάρη έχουν **μονάδες**: δεν συγκρίνουμε απευθείας w για %C και w για °C. Για σύγκριση κανονικοποιούμε τα χαρακτηριστικά.

Αν δύο χαρακτηριστικά είναι **συσχετισμένα** μεταξύ τους (π.χ. %C και θερμοκρασία βαφής), η ερμηνεία των βαρών γίνεται αναξιόπιστη (πολυσυγγραμμικότητα).

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πολλά χαρακτηριστικά: η δομή των δεδομένων

Στην πράξη κάθε δείγμα έχει **πολλά χαρακτηριστικά**. Τα δεδομένα οργανώνονται σε πίνακα: **I γραμμές** (δείγματα) × **J στήλες** (χαρακτηριστικά), όπως στο NumPy και το Pandas.

Δείγμα	%C	%Cr	T (°C)	t (min)	HV
M1	0,20	1,0	800	30	180
M2	0,40	1,5	850	45	235
M3	0,60	1,0	900	60	310
M4	0,35	1,8	850	30	260
M5	0,50	1,2	875	45	290

X: πίνακας χαρακτηριστικών (5 × 4) **y**

Κάθε **γραμμή** είναι ένα δείγμα x_i, κάθε **στήλη** ένα χαρακτηριστικό.

Ένα δείγμα = ένα διάνυσμα

$x_i = [x_{i1}, x_{i2}, \ldots, x_{ij}]$

Π.χ. M2 = [0,40 1,5 850 45]

Το ίδιο πλαίσιο με τη Διάλεξη 1

Ο πίνακας X και το διάνυσμα y είναι τα δεδομένα εκπαίδευσης κάθε επιβλεπόμενου μοντέλου.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Ένα βάρος για κάθε χαρακτηριστικό

Το μοντέλο παίρνει **ένα βάρος w_j για κάθε χαρακτηριστικό** και έναν σταθερό όρο b.

$$f(x) = w_1x_1 + w_2x_2 + \dots + w_Jx_J + b$$

Στο παράδειγμα των κραμάτων

$$HV = w_1 \cdot \%C + w_2 \cdot \%Cr + w_3 \cdot T + w_4 \cdot t + b$$

Κάθε w_j

πόσο αλλάζει η σκληρότητα όταν αλλάζει **μόνο** το χαρακτηριστικό j

Πρόσημο w_j

θετικό: αυξάνει τη σκληρότητα · αρνητικό: τη μειώνει

Γεωμετρικά

1 χαρακτηριστικό → ευθεία

2 → επίπεδο

J → υπερεπίπεδο

Η εκπαίδευση βρίσκει J + 1 παραμέτρους: όλα τα w_j και το b.

Συνάρτηση απώλειας και αναλυτική λύση

Ιδέα: ελαχιστοποιούμε το άθροισμα τετραγώνων.

Με πίνακες η λύση γράφεται σε **μία γραμμή**. $\hat{\mathbf{X}}$ είναι ο πίνακας X με μια επιπλέον στήλη από άσσους, για το b.

Απώλεια

$$l(\hat{w}) = \left(\hat{\mathbf{X}} \cdot \hat{w} - \vec{y}\right)^T \cdot \left(\hat{\mathbf{X}} \cdot \hat{w} - \vec{y}\right)$$

Λύση ελαχίστων τετραγώνων

$$\hat{\mathbf{w}}^* = \left(\hat{\boldsymbol{X}}^{\mathbf{T}} \cdot \hat{\boldsymbol{X}}\right)^{-1} \cdot \hat{\boldsymbol{X}}^{\mathbf{T}} \cdot \vec{\mathbf{y}}$$

Πότε υπάρχει η λύση;

- Όταν ο $\hat{X}^{\mathrm{T}}\hat{\mathrm{X}}$ αντιστρέφεται
- Χρειάζονται **περισσότερα δείγματα από παραμέτρους** (I ≥ J + 1)
- Τα χαρακτηριστικά να μην είναι γραμμικά εξαρτημένα

Πώς φτιάχνεται το $\hat{\mathrm{X}}$

$\hat{\mathrm{X}}$ (δείγματα × (J + 1))					$\vec{y}$
1	%C	%Cr	T (°C)	t (min)	HV
1	0,20	1,0	800	30	180
1	0,40	1,5	850	45	235
1	0,60	1,0	900	60	310
…	…	…	…	…	…

- Η πορτοκαλί στήλη με τους **άσσους** είναι για το **b**.
- Η σκληρότητα (HV) δεν μπαίνει στο $\hat{\mathrm{X}}$: είναι το $\vec{y}$.
- Εδώ J = 4 χαρακτηριστικά → J + 1 = 5 στήλες.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πόσο καλό είναι το μοντέλο;

Τρεις συνήθεις μετρικές για την παλινδρόμηση. Θα τις δούμε αναλυτικά στη διάλεξη Αξιολόγησης

MSE

Μέσο τετραγωνικό σφάλμα

$$MSE = (1/l)\cdot\Sigma(y_i - f(x_i))^2$$

Μονάδες: τετράγωνο της y

Όσο μικρότερο, τόσο καλύτερα

RMSE

Ρίζα του MSE

$$RMSE = \sqrt{MSE}$$

Στις ίδιες μονάδες με την y

π.χ. RMSE = 12 HV →
λάθος περίπου ±12 HV

R^2

Συντελεστής προσδιορισμού

$$R^2 = 1 - \Sigma(y_i - f_i)^2 / \Sigma(y_i - \bar{y})^2$$

1 = τέλειο, 0 = όσο ο μέσος όρος

π.χ. $R^2 = 0,9$ → εξηγεί το 90% της μεταβολής

Προσοχή: μετράμε την απόδοση σε **δεδομένα ελέγχου**, όχι μόνο σε αυτά της εκπαίδευσης.
Ένα μοντέλο με $R^2 = 1$ στα δεδομένα εκπαίδευσης μπορεί να γενικεύει πολύ άσχημα, όπως θα δούμε αμέσως μετά.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.



Ενότητα 4

Κανονικοποίηση Ridge και Lasso

Όταν τα δείγματα είναι λίγα

🎯 Με **λίγα δείγματα** (I) και **πολλά χαρακτηριστικά** (J), το βασικό γραμμικό μοντέλο υπερπροσαρμόζεται εύκολα.

Παράδειγμα: 5 κράματα, 4 χαρακτηριστικά + b = 5 άγνωστοι

$$HV = 40 \cdot \%C - 20 \cdot \%Cr + 1{,}8 \cdot T - 2{,}2 \cdot t - 1182$$

Κράμα	HV μετρημένη	HV μοντέλου	Σφάλμα
M1	180	180	0
M2	235	235	0
M3	310	310	0
M4	260	260	0
M5	290	290	0

Μηδενικό σφάλμα… αλλά;

- Το %Cr εμφανίζεται να **μειώνει** τη σκληρότητα
- Το b = −1182 δεν έχει φυσικό νόημα
- Σε νέο κράμα οι προβλέψεις είναι αναξιόπιστες

Δύο λύσεις

1. Λιγότερα χαρακτηριστικά: κρατάμε μόνο τα σημαντικά

2. Κανονικοποίηση (regularization)

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Η ιδέα της κανονικοποίησης

Μοντέλα με **υψηλή ακρίβεια** και ταυτόχρονα **χαμηλή πολυπλοκότητα**.
Προσθέτουμε στην απώλεια έναν όρο που «τιμωρεί» τα μεγάλα βάρη.

Σφάλμα
Πόσο καλά ταιριάζει στα δεδομένα

$+$ **λ** $\times$

Ποινή
Πόσο μεγάλα είναι τα βάρη

$=$ **νέα απώλεια**

Ο ρόλος του λ

λ = 0

Καμία ποινή: το απλό γραμμικό μοντέλο (OLS).

λ κατάλληλο

Ισορροπία ανάμεσα σε ακρίβεια και απλότητα: καλύτερη γενίκευση.

λ πολύ μεγάλο

Τα βάρη πλησιάζουν το 0: το μοντέλο γίνεται υπερβολικά απλό (υποπροσαρμογή).

Το λ είναι **υπερπαράμετρος**: δεν μαθαίνεται από την εκπαίδευση, το επιλέγουμε εμείς (συνήθως με διασταυρωμένη επικύρωση).

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Ridge: κανονικοποίηση L2

Η Ridge προσθέτει ως ποινή το **άθροισμα των τετραγώνων** των βαρών.

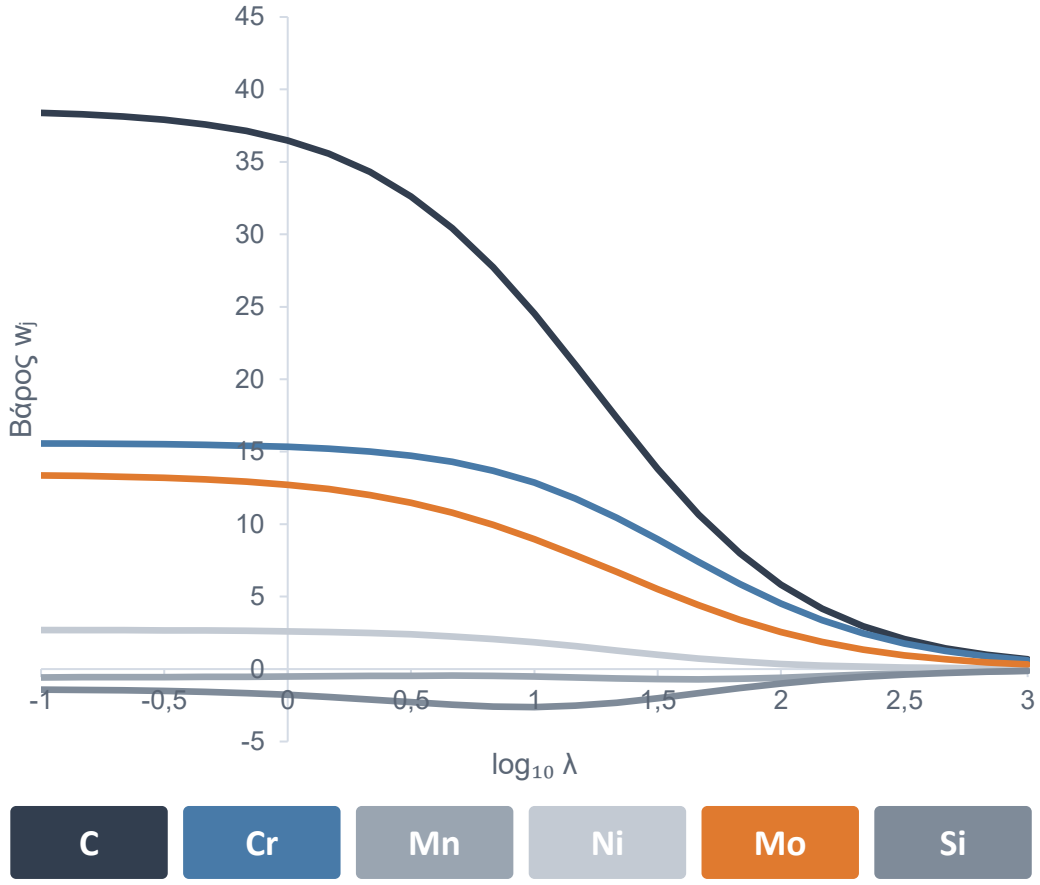
$$\boldsymbol{l}(\hat{\boldsymbol{w}}) = (\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{w}} - \vec{\boldsymbol{y}})^{\boldsymbol{T}} \cdot (\hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{w}} - \vec{\boldsymbol{y}}) + \boldsymbol{\lambda}\hat{\boldsymbol{\omega}}^{\boldsymbol{T}} \cdot \hat{\boldsymbol{w}}$$

Έχει και αυτή **αναλυτική λύση**:

$$\hat{\mathbf{w}}^{*} = \left(\hat{\boldsymbol{X}}^{\mathbf{T}} \cdot \hat{\mathbf{x}} + \boldsymbol{\lambda}\hat{\mathbf{I}}\right)^{-\mathbf{1}} \cdot \hat{\boldsymbol{X}}^{\mathbf{T}} \cdot \vec{\mathbf{y}}$$

Τι κάνει

- **«Συρρικνώνει»** όλα τα βάρη προς το 0, αλλά σπάνια τα μηδενίζει
- Σταθεροποιεί τη λύση όταν τα χαρακτηριστικά είναι **συσχετισμένα**
- Το λI κάνει τον πίνακα πάντα αντιστρέψιμο (για λ > 0)



Συνθετικά δεδομένα για επίδειξη: 30 κράματα, 6 στοιχεία. Η σκληρότητα εξαρτάται μόνο από τα C, Cr, Mo.

Lasso: κανονικοποίηση L1

Η Lasso βάζει ως ποινή το **άθροισμα των απόλυτων τιμών** των βαρών.

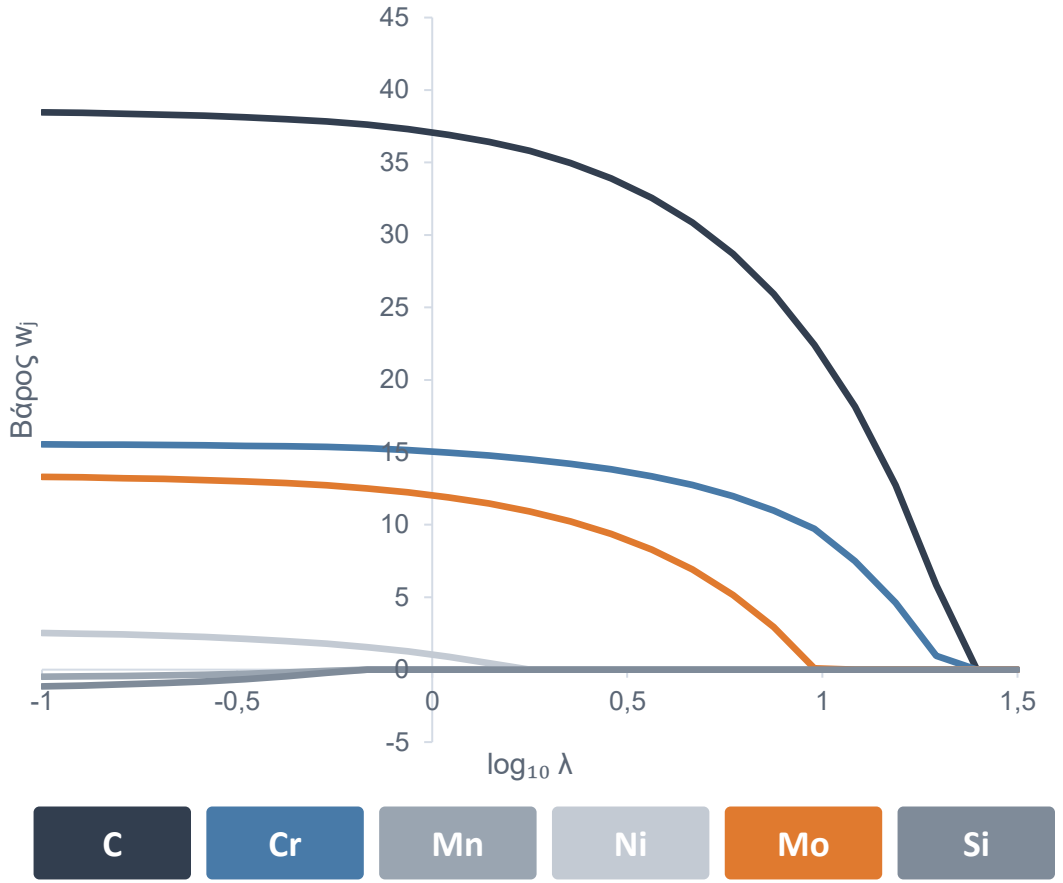
$$\mathbf{l}(\hat{\boldsymbol{\omega}}) = (\hat{\mathbf{x}} \cdot \hat{\mathbf{w}} - \vec{\mathbf{y}})^{\mathbf{T}} \cdot (\hat{\mathbf{x}} \cdot \hat{\mathbf{w}} - \vec{\mathbf{y}}) + \boldsymbol{\lambda} \sum_{\mathbf{j=1}}^{\mathbf{J}} \|\boldsymbol{\omega}_{\mathbf{j}}\|$$

Δεν έχει αναλυτική λύση

Η απόλυτη τιμή δεν παραγωγίζεται στο 0. Τη λύνουμε με **αριθμητικό solver** (π.χ. από το SciPy).

Μηδενίζει βάρη

Καθώς αυξάνεται το λ, βάρη γίνονται **ακριβώς 0**. Άρα η Lasso κάνει ταυτόχρονα και **επιλογή χαρακτηριστικών**: κρατά μόνο τα σημαντικά.



Ίδια δεδομένα: τα Mn, Ni, Si μηδενίζονται γρήγορα. Τα C, Cr, Mo μένουν.

Ridge ή Lasso;

	Ridge (L2)	Lasso (L1)
Ποινή	$\lambda \cdot \Sigma\, w_j^2$	$\lambda \cdot \Sigma\, \lvert w_j \rvert$
Αναλυτική λύση	Ναι	Όχι (solver)
Επίδραση στα βάρη	Τα μικραίνει όλα	Μηδενίζει κάποια
Επιλογή χαρακτηριστικών	Όχι	Ναι
Πότε τη χρησιμοποιούμε	Πολλά χρήσιμα, συσχετισμένα χαρακτηριστικά	Λίγα από τα πολλά χαρακτηριστικά μετράνε

Elastic Net

Συνδυάζει τις δύο ποινές (L1 + L2)

Επιλογή του λ

Δοκιμάζουμε πολλές τιμές και κρατάμε αυτή με το μικρότερο σφάλμα στη **διασταυρωμένη επικύρωση**.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.



Ενότητα 5

Λογιστική Παλινδρόμηση και Ταξινόμηση

Τι είναι η ταξινόμηση;

Στην **ταξινόμηση** (classification) η έξοδος δεν είναι αριθμός, αλλά **κατηγορία**. Το μοντέλο αποφασίζει σε ποια ομάδα ανήκει ένα δείγμα.

Παλινδρόμηση

Πόσο;

Πόση θα είναι η σκληρότητα;

→ 268 HV

Ταξινόμηση

Ποιο;

Είναι το υλικό κυτταροτοξικό;

→ ναι / όχι

Δύο είδη

Δυαδική

2 κατηγορίες (1 / 0)

Πολλών κατηγοριών

3 ή περισσότερες

Ερώτημα στα υλικά	Κατηγορίες
Είναι το υλικό κυτταροτοξικό;	ναι / όχι
Θα αστοχήσει το εμφύτευμα σε 5 έτη;	ναι / όχι
Τι ιστός αναπτύσσεται γύρω του;	οστό / ινώδης
Ποια φάση δείχνει η μικρογραφία;	φερρίτης / περλίτης / μαρτενσίτης

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Λογιστική παλινδρόμηση: τι είναι και γιατί

Η λογιστική παλινδρόμηση είναι ένα **γραμμικό μοντέλο ταξινόμησης**: υπολογίζει την **πιθανότητα** ένα δείγμα να ανήκει στην κατηγορία 1 και αποφασίζει με βάση αυτήν.

1. Χαρακτηριστικά

απελευθέρωση ιόντων $x_1 = 4$

τραχύτητα $x_2 = 0,8$

→

2. Γραμμικός συνδυασμός

$y = 0,8 \cdot 4 + 0,5 \cdot 0,8 - 3$

$= 0,6$

→

3. Λογιστική συνάρτηση

e^(−0,6) ≈ 0,55

z = 1 / (1 + 0,55) = 0,65

→

4. Απόφαση

0,65 ≥ 0,5

→ κυτταροτοξικό

Γιατί «παλινδρόμηση»;

Γιατί μέσα του υπολογίζει έναν **συνεχή** αριθμό ($y = w^T x + b$), όπως η παλινδρόμηση. Η κατηγορία προκύπτει **στο τέλος**, από την πιθανότητα.

Γιατί τη χρησιμοποιούμε

- Απλή και γρήγορη
- Δίνει **πιθανότητα**, όχι μόνο ναι/όχι
- Το πρόσημο κάθε w δείχνει αν το χαρακτηριστικό **αυξάνει** τον κίνδυνο

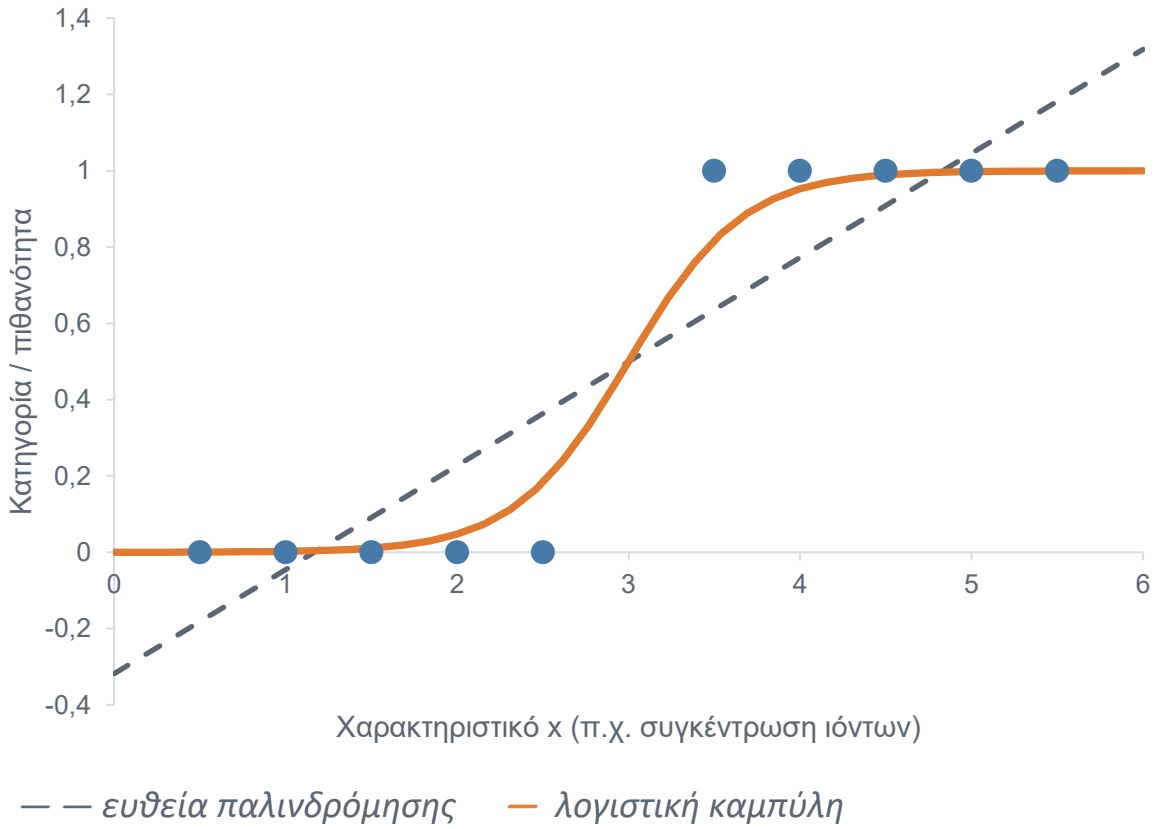
Πότε δεν αρκεί

Όταν οι κατηγορίες **δεν χωρίζονται με ευθεία**. Τότε χρειαζόμαστε πυρήνες ή μη γραμμικά μοντέλα (δέντρα απόφασης).

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Από την παλινδρόμηση στην ταξινόμηση

Τα γραμμικά μοντέλα σχεδιάστηκαν για **παλινδρόμηση**.
Στην ταξινόμηση η έξοδος είναι **κατηγορία**, π.χ. κυτταροτοξικό (1) ή όχι (0).



Γιατί όχι απλή ευθεία;

- Δίνει τιμές **κάτω από 0 και πάνω από 1**
- Δεν ερμηνεύεται ως πιθανότητα

Η λύση

Περνάμε την έξοδο του γραμμικού μοντέλου από μια **ομαλή συνάρτηση** που τη «συμπιέζει» στο [0, 1]: τη **λογιστική συνάρτηση**.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Η λογιστική (σιγμοειδής) συνάρτηση

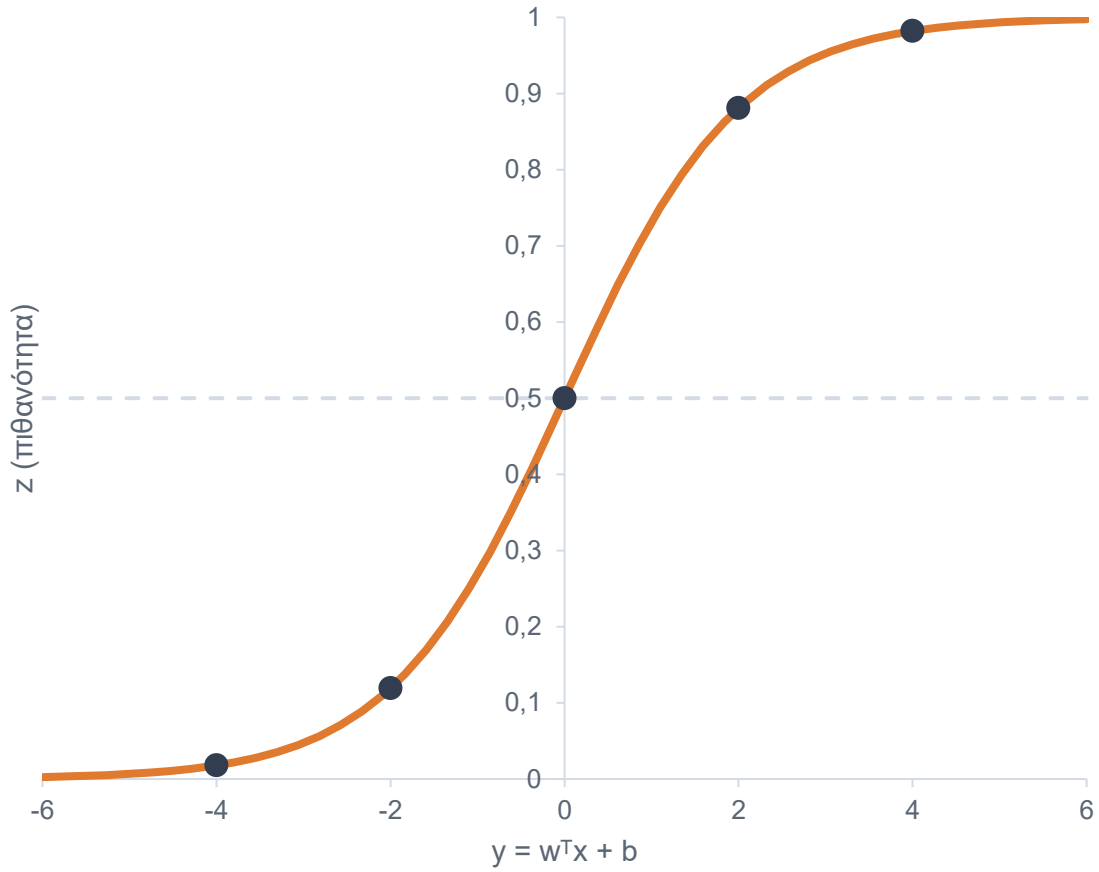
Η έξοδος του γραμμικού μοντέλου $y = w^T x + b$ γίνεται **πιθανότητα** z ανάμεσα σε 0 και 1

$$z = 1 / (1 + e^{-y})$$

$y = w^T x + b$	z	Ερμηνεία
−4	0,02	σχεδόν σίγουρα 0
−2	0,12	μάλλον 0
0	0,50	αβέβαιο
+2	0,88	μάλλον 1
+4	0,98	σχεδόν σίγουρα 1

P(κατηγορία 1) = z, P(κατηγορία 0) = 1 − z

Λόγος πιθανοτήτων (odds): $z / (1 − z) = e^y$



Κανόνας απόφασης και όριο απόφασης

Επιλέγουμε την κατηγορία με τη **μεγαλύτερη πιθανότητα**: αν z ≥ 0,5 → κατηγορία 1, αλλιώς 0.

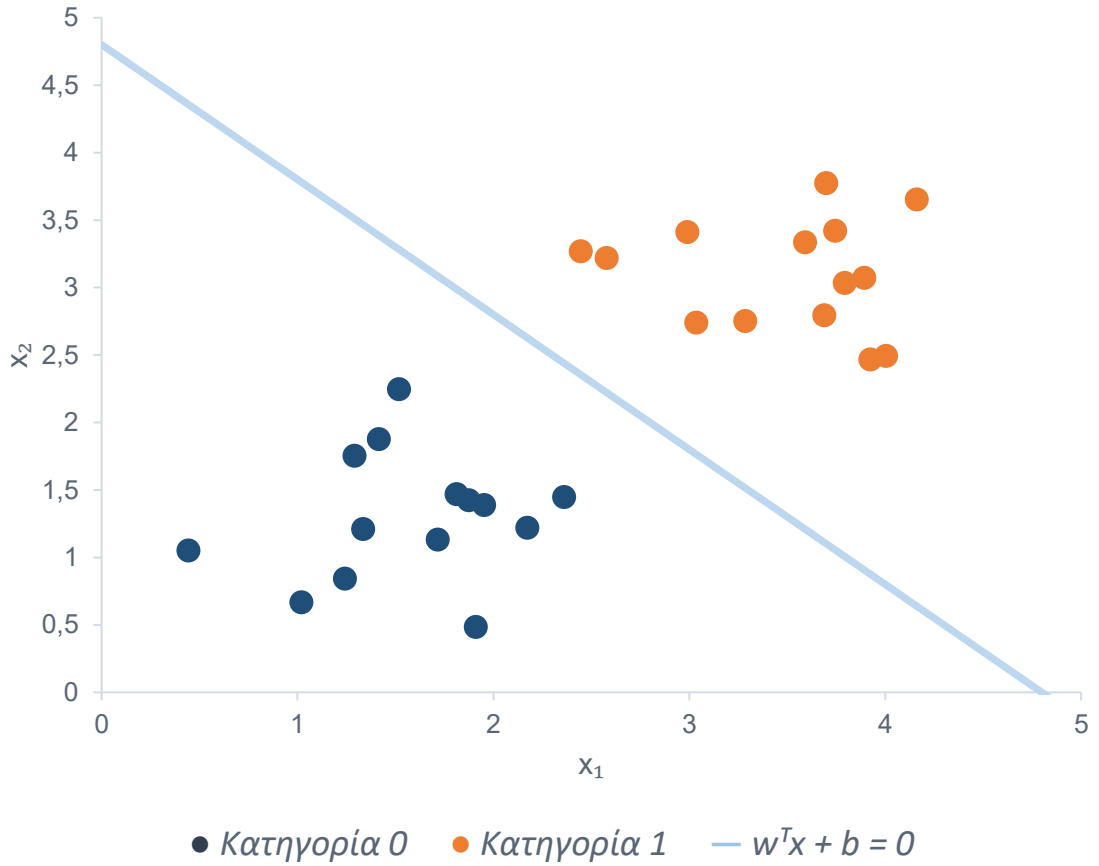
Γιατί το όριο είναι γραμμικό

z = 0,5 ακριβώς όταν $\mathbf{w}^T\mathbf{x} + \mathbf{b} = \mathbf{0}$. Αυτό είναι ευθεία (2 χαρακτηριστικά) ή υπερεπίπεδο (J χαρακτηριστικά).

Το κατώφλι 0,5 δεν είναι υποχρεωτικό

Το βιβλίο σημειώνει ότι μπορεί να επιλεγεί άλλο κατώφλι.

Π.χ. για ένα εμφύτευμα, αν το λάθος «ασφαλές ενώ είναι κυτταροτοξικό» κοστίζει πολύ, χαμηλώνουμε το κατώφλι (π.χ. 0,3).



Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Απώλεια: διασταυρούμενη εντροπία

Στην ταξινόμηση δεν μετράμε αποστάσεις, αλλά **πόσο πιθανή** έκρινε το μοντέλο τη σωστή κατηγορία. Η απώλεια λέγεται διασταυρούμενη εντροπία (cross-entropy).

$$L = -\Sigma \left[y_i \cdot \ln P_1 + (1 - y_i) \cdot \ln P_0 \right]$$

Η διαίσθηση: ένα δείγμα με σωστή κατηγορία 1

Μοντέλο: $P_1 = 0,9$

Απώλεια = $-\ln 0,9$ = **0,11**

σωστό και σίγουρο

Μοντέλο: $P_1 = 0,6$

Απώλεια = $-\ln 0,6$ = **0,51**

σωστό αλλά αβέβαιο

Μοντέλο: $P_1 = 0,1$

Απώλεια = $-\ln 0,1$ = **2,30**

λάθος και σίγουρο

Τα λάθη με «σιγουριά» τιμωρούνται πολύ. Η εκπαίδευση βρίσκει τα w, b που ελαχιστοποιούν τη συνολική απώλεια.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Πολλές κατηγορίες: Softmax και one-hot

Για K > 2 κατηγορίες, το μοντέλο βγάζει **K τιμές**, μία για κάθε κατηγορία, και η **Softmax** τις μετατρέπει σε πιθανότητες που αθροίζουν 1.

$$P_k = \exp(y_k) / \Sigma_j \exp(y_j)$$

Παράδειγμα: τύπος φάσης σε 3 κατηγορίες

	Φερρίτης	Περλίτης	Μαρτενσίτης
Τιμή y_k	2,0	1,0	0,1
Softmax P_k	0,66	0,24	0,10
Πρόβλεψη (one-hot)	1	0	0

Ενδεικτικές τιμές.

Κωδικοποίηση one-hot

Η κατηγορία γράφεται ως διάνυσμα με **ένα 1** στη θέση της και 0 αλλού: φερρίτης = [1, 0, 0].

Τι αλλάζει στο μοντέλο

Τα βάρη γίνονται **πίνακας (J+1) × K**: μία στήλη βαρών για κάθε κατηγορία.

Εναλλακτικά: πολλοί δυαδικοί ταξινομητές (μία κατηγορία έναντι των υπολοίπων).

Πόσο καλός είναι ο ταξινομητής;

Η πιο απλή μετρική είναι η **ακρίβεια** (accuracy).
Ο **πίνακας σύγχυσης** δείχνει όμως και **τι είδους** λάθη κάνει το μοντέλο.

Παράδειγμα: 100 δείγματα βιοϋλικών

Πρόβλεψη μοντέλου

Πραγματικό	Κυτταροτοξικό	Ασφαλές
Κυτταροτοξικό	18	7
Ασφαλές	5	70

Ακρίβεια = (18 + 70) / 100 = **0,88**

Ενδεικτικές τιμές.

Τι κρύβει η ακρίβεια

7 κυτταροτοξικά υλικά χαρακτηρίστηκαν **ασφαλή**. Αυτό το λάθος είναι πολύ σοβαρότερο από το αντίθετο.

Στη διάλεξη Αξιολόγησης

Ευαισθησία, ειδικότητα, precision, F1, καμπύλη ROC: μετρικές που ξεχωρίζουν τα είδη των λαθών.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.



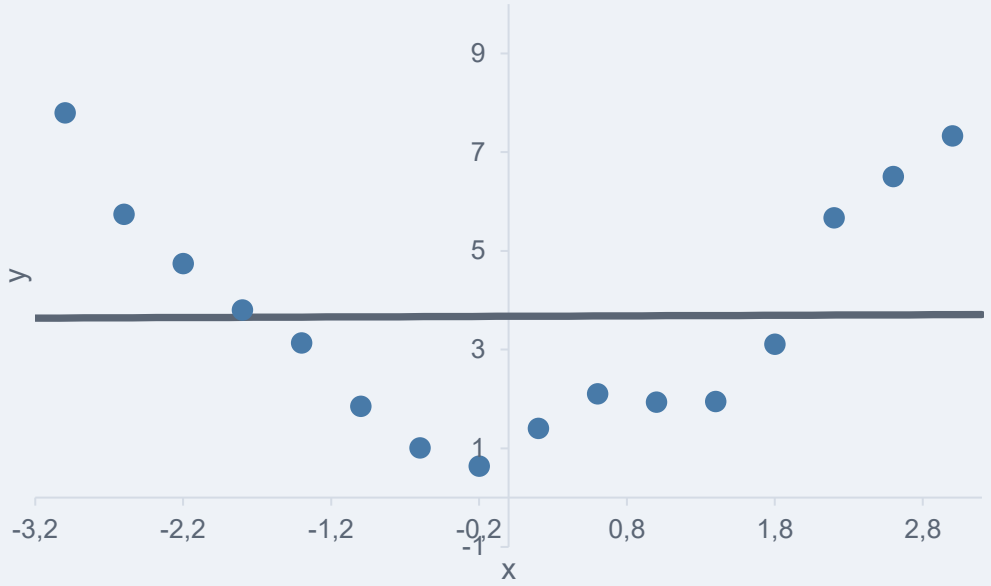
Ενότητα 6

Μη Γραμμικά Μοντέλα μέσω Συναρτήσεων Πυρήνα

Το όριο των γραμμικών μοντέλων

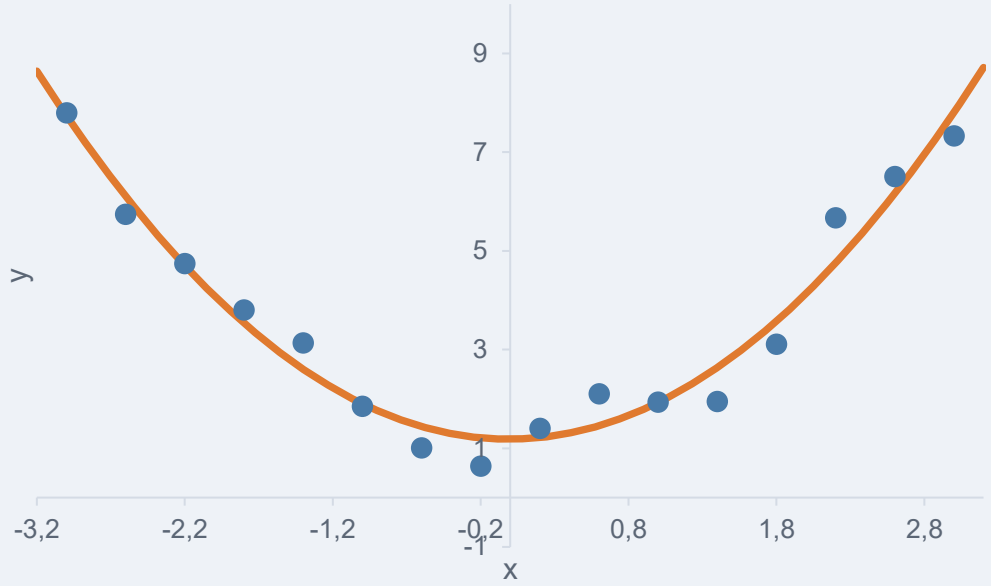
Όταν τα δεδομένα **δεν ακολουθούν ευθεία**, ένα γραμμικό μοντέλο δεν μπορεί να τα περιγράψει ικανοποιητικά, όσο καλά κι αν το εκπαιδεύσουμε.

Ευθεία: f(x) = w·x + b



Μεγάλα, συστηματικά υπόλοιπα

Με x^2: $f(x) = w_1x + w_2x^2 + b$



Ακολουθεί τα δεδομένα

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Απεικόνιση σε περισσότερες διαστάσεις

Το κόλπο: φτιάχνουμε **νέα χαρακτηριστικά**, π.χ. x → [x, x^2]. Στον νέο χώρο το πρόβλημα γίνεται **γραμμικό**.

$$f(x) = w_1 \cdot x + w_2 \cdot x^2 + b$$

= γραμμικό μοντέλο στις μεταβλητές $t_1 = x$ και $t_2 = x^2$

Απεικόνιση ϕ(x)

Η συνάρτηση που «τεντώνει» τα δεδομένα σε περισσότερες διαστάσεις (stretching function).

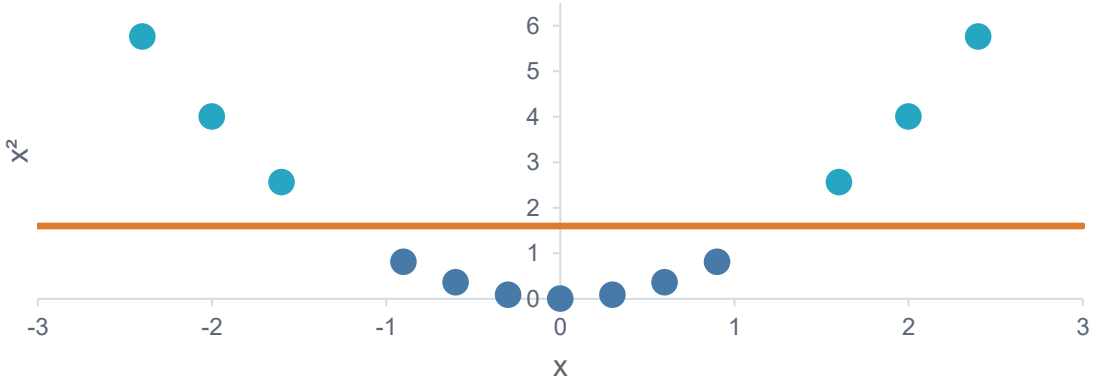
Προσοχή

Μια **γραμμική** απεικόνιση (π.χ. $x_1 + x_2$) δεν βοηθά: τα δεδομένα μένουν γραμμικά.

1 διάσταση: καμία ευθεία δεν τα χωρίζει



2 διαστάσεις [x, x^2]: τα χωρίζει μια ευθεία



Το τέχνασμα του πυρήνα (kernel trick)

Η απεικόνιση φ μπορεί να είναι **ακριβή** (πολλές νέες διαστάσεις) και **δύσκολο να επιλεγεί**.
Η συνάρτηση πυρήνα μας επιτρέπει να την παρακάμψουμε.

Συνάρτηση πυρήνα

Παίρνει **δύο δείγματα** και δίνει **έναν αριθμό**: πόσο «όμοια» είναι στον νέο χώρο.

$$k(a, b) = \phi(a)^T \cdot \phi(b)$$

Υπολογίζεται χωρίς να φτιάξουμε ποτέ το φ

Πώς γίνεται η πρόβλεψη

Για ένα νέο δείγμα x, το μοντέλο συγκρίνει το x με **κάθε δείγμα εκπαίδευσης** x_i και αθροίζει με βάρη z_i:

$$f(x) = \Sigma\, z_i \cdot k(x, x_i)$$

Η διαίσθηση: η πρόβλεψη για ένα νέο κράμα είναι ένας σταθμισμένος συνδυασμός των κραμάτων εκπαίδευσης, με μεγαλύτερο βάρος σε όσα του **μοιάζουν** περισσότερο.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

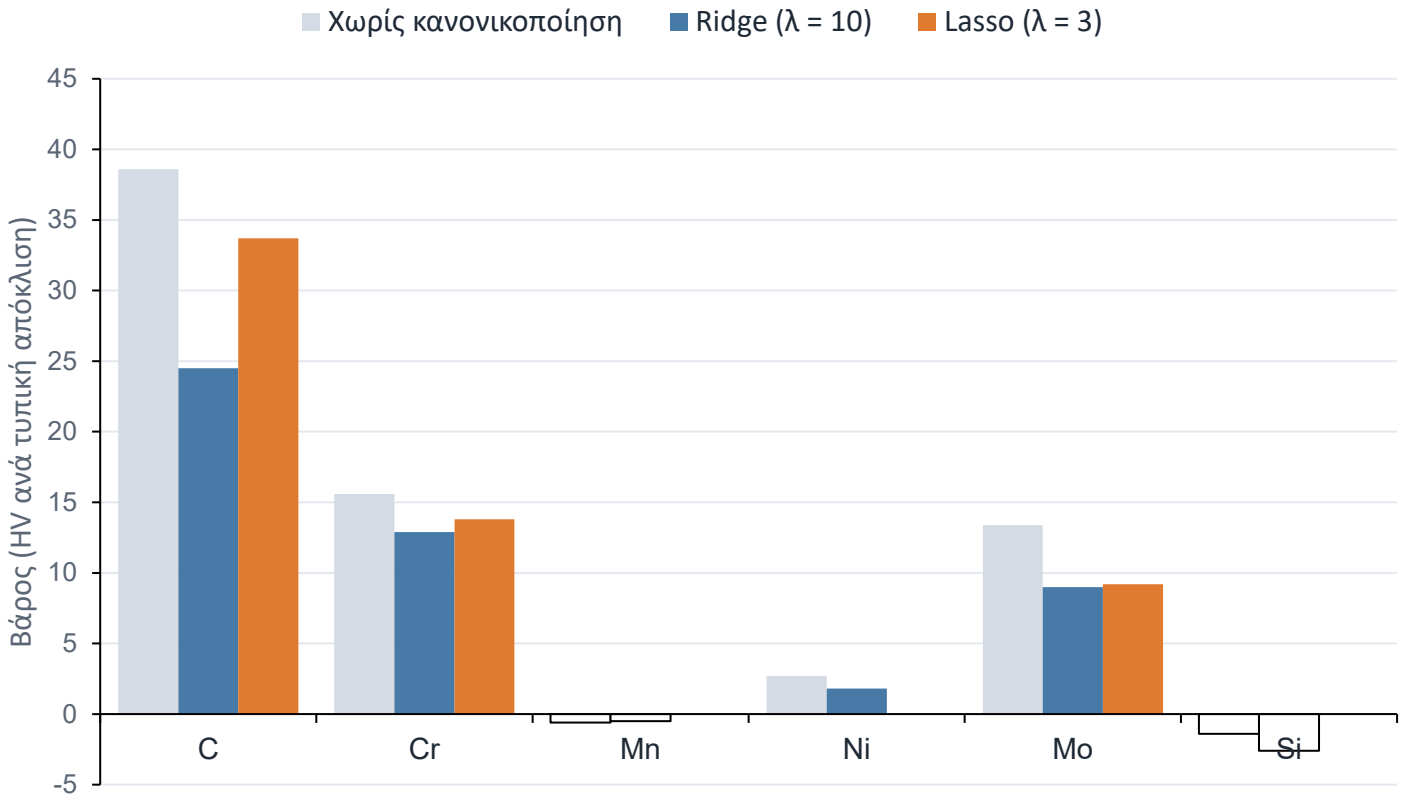


Ενότητα 7

Εφαρμογές στην Επιστήμη Υλικών και Σύνοψη

Εφαρμογή: σημαντικά κραματικά στοιχεία

Με πολλά κραματικά στοιχεία και λίγα δείγματα, η Lasso δείχνει **ποια στοιχεία μετράνε**.
Το αποτέλεσμα είναι ερμηνεύσιμο, κάτι πολύτιμο στον σχεδιασμό υλικών.



Τα χαρακτηριστικά είναι κανονικοποιημένα, ώστε τα βάρη να συγκρίνονται μεταξύ τους.

Τι βλέπουμε

- Η **Lasso** μηδενίζει Mn, Ni, Si
- Η **Ridge** τα μικραίνει, αλλά τα κρατά

Πρακτικό όφελος

Λιγότερα στοιχεία προς μελέτη, πιο **στοχευμένα πειράματα** και απλούστερο μοντέλο.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Σύνοψη των γραμμικών μοντέλων

Το μοντέλο το διαλέγουμε από το **είδος της εξόδου** και από το **πόσα δεδομένα** έχουμε.

Ερώτηση για το πρόβλημα	Μοντέλο	Παράδειγμα υλικών
Συνεχής έξοδος, ένα χαρακτηριστικό	Απλή γραμμική	Σκληρότητα από το %C
Συνεχής έξοδος, πολλά χαρακτηριστικά	Πολλαπλή	Σκληρότητα από %C, %Cr, θερμοκρασία, χρόνο
Έξοδος κατηγορία	Λογιστική	Κυτταροτοξικό ή όχι
Πολλά χαρακτηριστικά, λίγα δείγματα	Ridge / Lasso	Ποια κραματικά στοιχεία μετράνε
Η σχέση δεν είναι ευθεία	Με πυρήνα	Ιδιότητα που ανεβαίνει και μετά πέφτει με τη θερμοκρασία

Κοινός πυρήνας όλων: ο γραμμικός συνδυασμός **w·x + b**. Αλλάζει μόνο η απώλεια, η ποινή ή η συνάρτηση που εφαρμόζεται στην έξοδο.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.

Βασικά συμπεράσματα

1. Το γραμμικό μοντέλο προβλέπει με **γραμμικό συνδυασμό** των χαρακτηριστικών: $f(x) = w^T x + b$
2. Τα w, b βρίσκονται με **ελάχιστα τετράγωνα**: με αναλυτική λύση ή αριθμητικά (solver)
3. Με λίγα δείγματα, η **κανονικοποίηση** (Ridge, Lasso) περιορίζει την υπερπροσαρμογή
4. Η **λογιστική παλινδρόμηση** μετατρέπει το γραμμικό μοντέλο σε ταξινομητή με πιθανότητες
5. Οι **πυρήνες** δίνουν μη γραμμικές δυνατότητες, με κόστος σε μνήμη και υπολογισμό

Στο επόμενο μάθημα

Δέντρα Απόφασης

- Μη γραμμικά όρια απόφασης
- Κανόνες «αν… τότε…»
- Εντροπία και δείκτης Gini

Τα γραμμικά μοντέλα είναι το **πρώτο μοντέλο που δοκιμάζουμε** σε ένα νέο πρόβλημα: γρήγορα, ερμηνεύσιμα και σημείο αναφοράς για όλα τα υπόλοιπα.

Βασική Πηγή: Liu, Z. (2025). *Artificial Intelligence for Engineers*.