

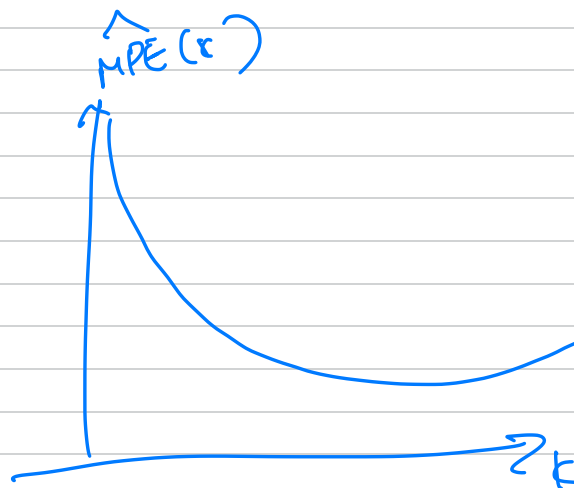
14-3-2025

Exempm

$MPE(k)$

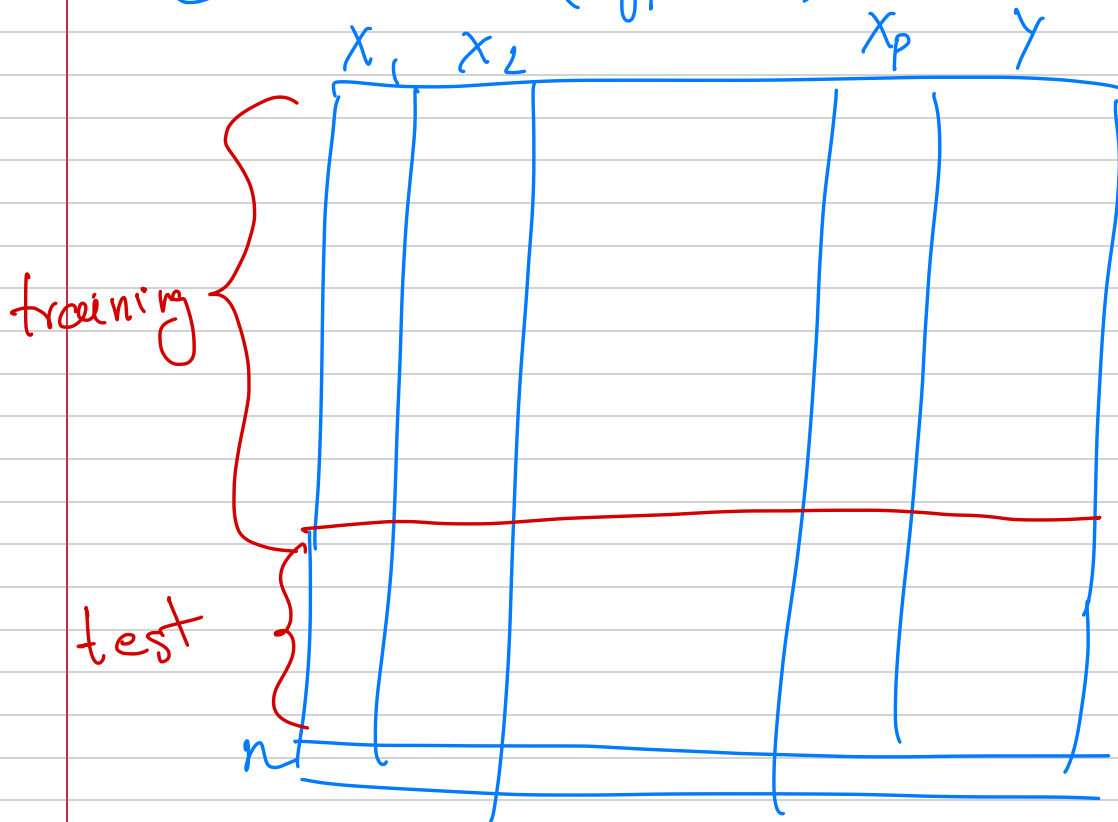
Train	$k=1$	$k=5$	$k=10$	$k=20$	$k=100$
1	92.00				
2					
.					
.					
.					
N					

mean
($k=1$)



Τι γίνεται όταν το training set προέρχεται από παρατηρησιμότητες;

Dataset (παρατηρησιμότητες)



Προέγγιση Χωρίζουμε το data set

σε n_1, n_2 παρατηρήσεις ($n_1 + n_2 = n$)

↓ training ↓ testing

Εφαρμόζουμε αλγ. 1 σε training $\Rightarrow \hat{f}_1(x)$
αλγ. 2 " " " $\Rightarrow \hat{f}_2(x)$

$$\left. \begin{aligned} \hat{y}_1 &= \hat{f}_1(x) & x \in \text{test} \\ \hat{y}_2 &= \hat{f}_2(x), & x \in \text{test} \end{aligned} \right\} \text{Σφάλμα} \sim \left(\underbrace{\hat{y}_1 - y_{\text{test}}}_{(\hat{y}_2 - y_{\text{test}})} \right)$$

Παραμένει το πρόβλημα ότι
έχω μόνο ένα training set

Αντίστοιχα στη στατιστική

Εάν ότι έχω ένα δείγμα $\sim n$ (αριθμός)
και πληθυσμό, $X \sim F$

$$\Delta E \quad \mu = E(X) \quad \left\{ \begin{array}{l} \hat{\mu} \\ L(x) \\ u(x) \end{array} \right\} \quad \left. \begin{array}{l} \text{ομαρ.} \\ \text{συνδυασμός} \end{array} \right\}$$

$\Delta E = (L, u) \leftarrow$ τιμές μεταβλητή

? $p = P(L \leq \mu \leq u) \leftarrow$

L, u τιμές μεταβλητή

που προέρχονται από $X \sim F$ (άγνωστη)

$\rightarrow p = \%$ διαστήματος περιόδου n
για τα οποία αν υπολογίσω
τα L, u θα προκύψουν:

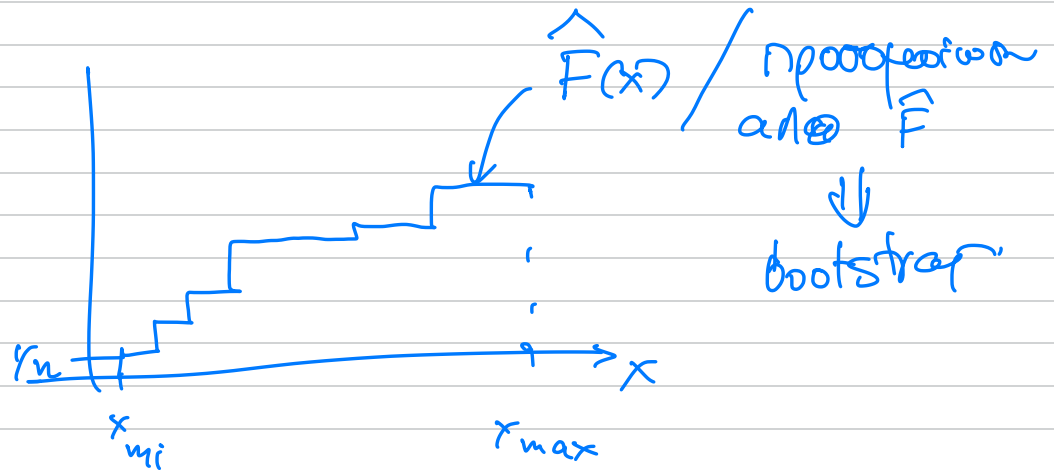
$$L \leq \mu \leq u$$

Resampling Methods

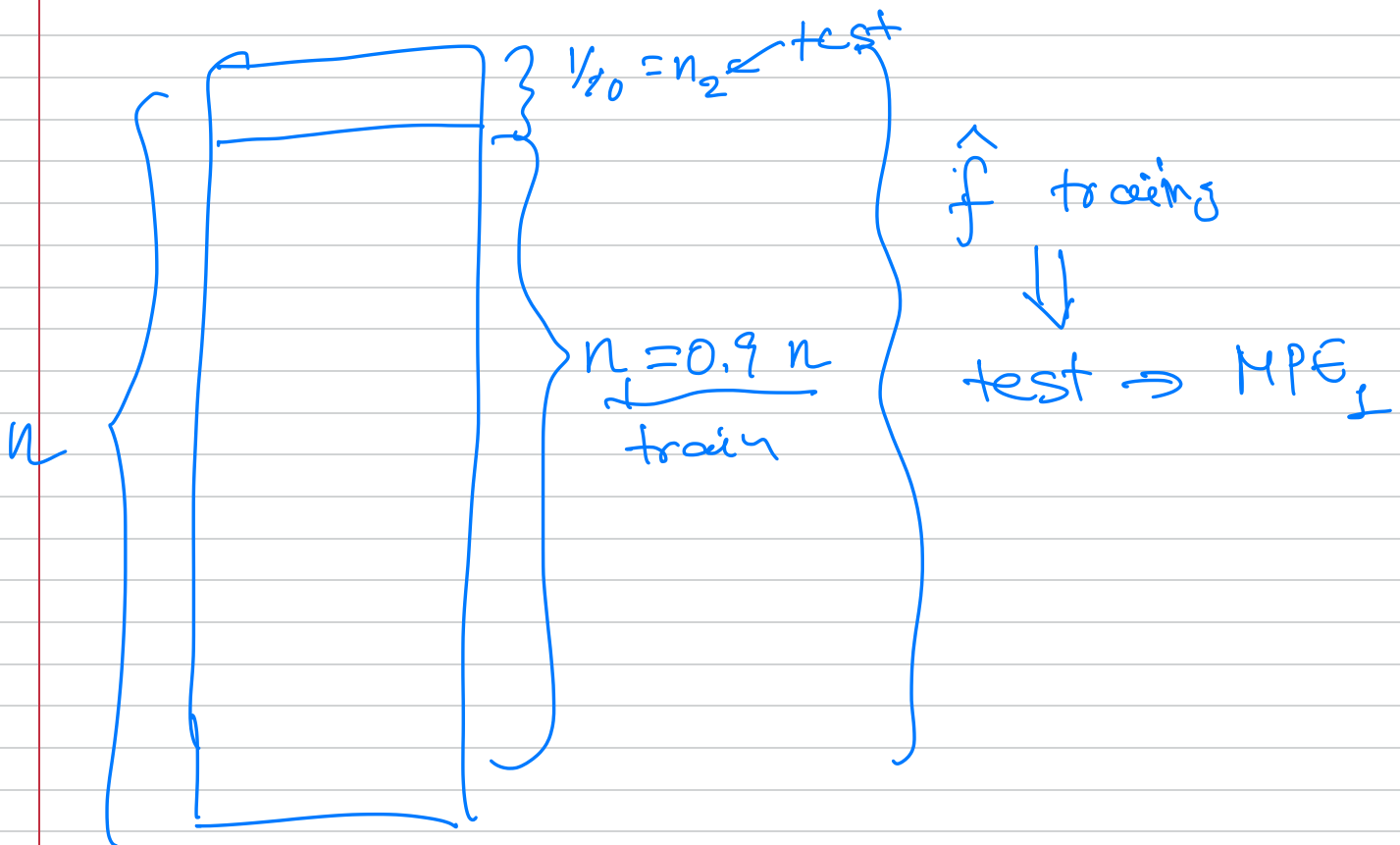
$\mathcal{D}^{\text{train}}$
(Bootstrap)
Cross Validation
 $\leftarrow \mathcal{D}^{\text{test}}$

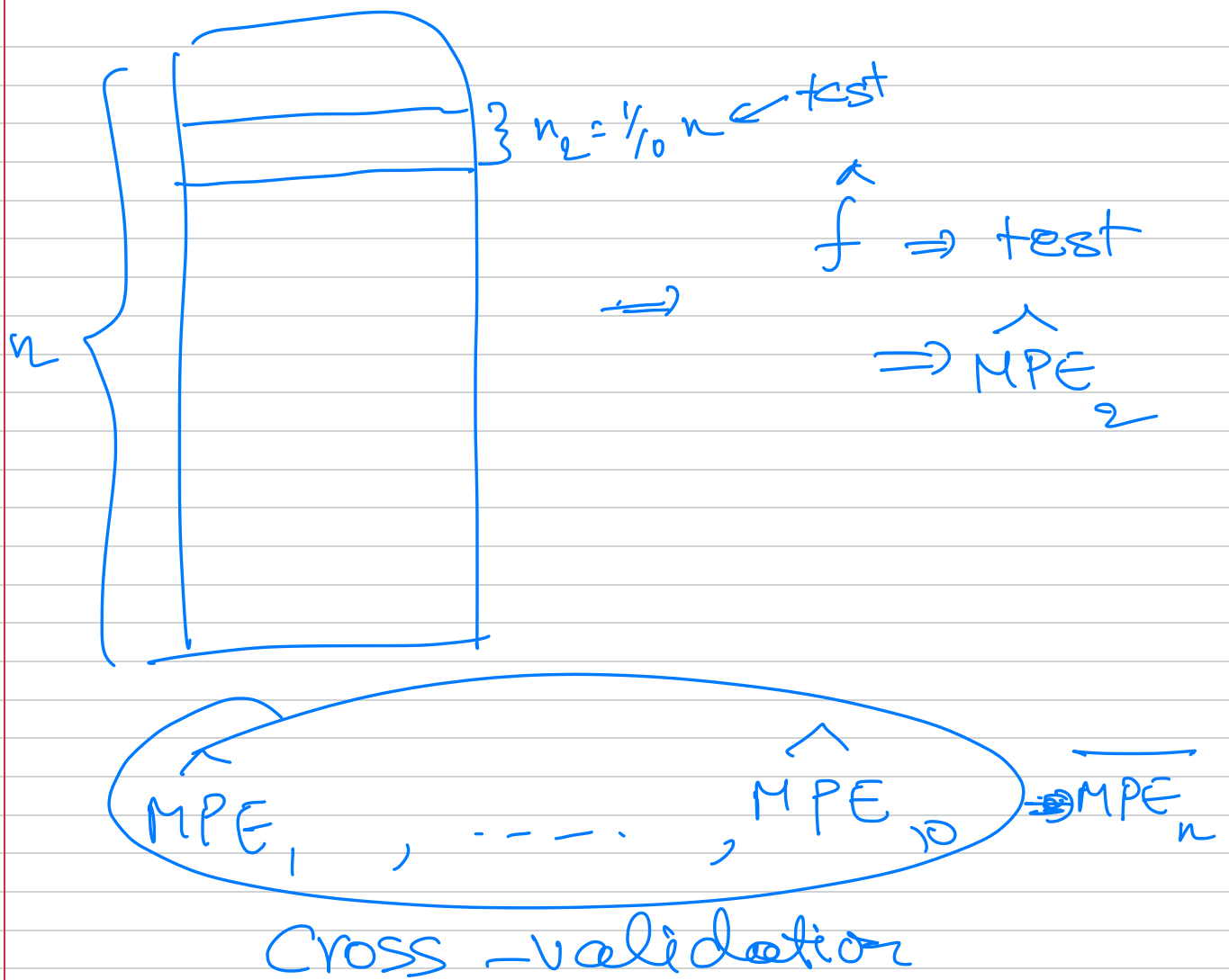
①

Deixe $\Rightarrow \hat{F}(x)$



②





Ανάλυση για Ποσοτική Y

(x, y)

Regression + ΕΛΕΚΤΑΧΕΣ

Στατιστικό μοντέλο (προσθετικό
additive model)

$$Y = f(x) + \varepsilon$$

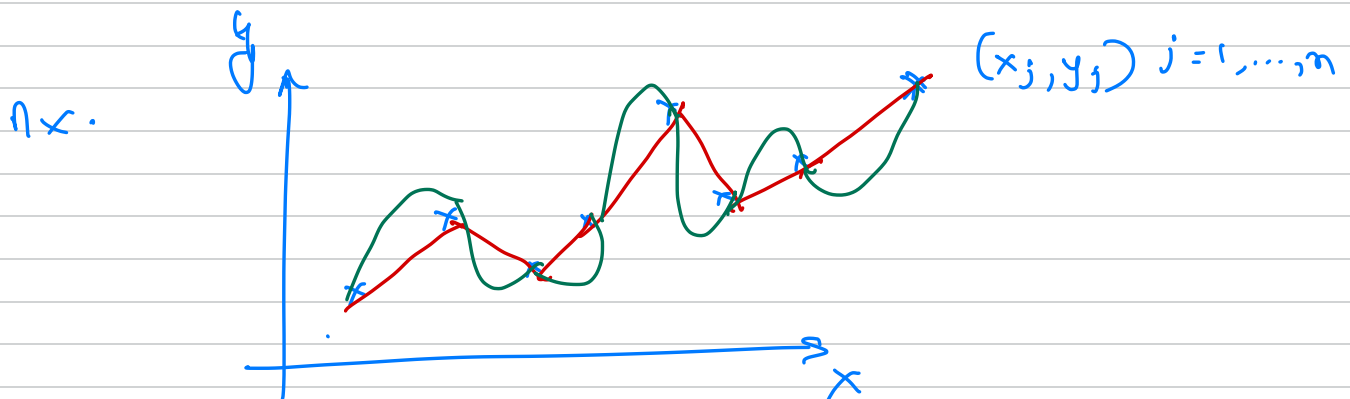
ε τυχόν
μεταβλητό
 $E(\varepsilon) = 0$

$$f(x) = E(Y|X=x)$$

$$X = (X_1, \dots, X_p)$$

ανάμεση
παιγνίσματος
(regression function)

Function approximation $\Rightarrow \hat{f}(x)$
(on training set)



Χωρίς υποθέσεις για την $f(x) \Rightarrow$ ζεζρημένο
(περιορισμοί
σε καμπύλες/εξισώσεις) πρόβλημα

Κάνουμε υπόθεση για την f

$\Leftrightarrow$ ψάχνουμε την "καλύτερη" προσέγγιση
 $\hat{f}(\cdot)$ όπου $\hat{f} \in \mathcal{C}$

Παραδείγματα

1) Γραμμική Παφενδρόμηση

$$f(x) = b_0 + b_1 x_1 + \dots + b_p x_p \quad (\text{γραμμικό} \\ \text{μοτέλο})$$

2) $f(x) = b_0 + b_1 x_1 + b_2 x_1^2 + b_3 x_2^3 + b_4 x_1 x_2$

Γραμμικό μοτέλο (Γραμμικό ως προς b)

Πολυωνυμική παφενδρόμηση

3) $f(x) = b_0 + b_1 e^{x_1} + b_2 x_2 \log x_1$

γραμμικό μοτέλο

με γραμμική παφενδρόμηση

4) $f(x) = \alpha_0 e^{b_0 x_1} + \alpha_1 \log(x_1^2 + c)$

$(\alpha_0, b_0, \alpha_1, c)$ αγνωστές παράμετροι
με γραμμικό μοτέλο
" " παφενδρόμηση.

Προσεγγιστικές Παφινδρόμους

Κριτήριο : $RSS(f) = \sum_{i=1}^n (y_i - f(x_i))^2$

Γραμμικά Μοντέλα : $\min_f \{ RSS(f) : f(x) = b_0 + b_1 x_1 + \dots + b_p x_p \}$

$\Leftrightarrow \min_{(b_0, b_1, \dots, b_p)} \left\{ \sum_{i=1}^n (y_i - b_0 - b_1 x_{i1} - \dots - b_p x_{ip})^2 \right\}$

① Μέθοδοι Εξομάλυνσης (Regularized regression)

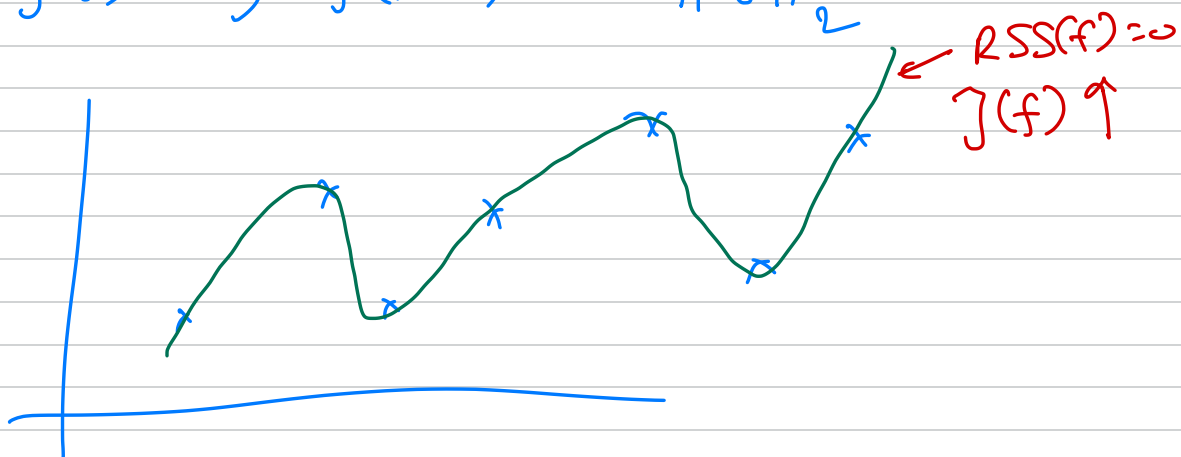
$\min_f \{ RSS(f) :$

$f(x) = b_0 + b_1 x_1 + \dots + b_p x_p$

και $J(f) \leq t \leftarrow \}$ (εξομάλυνση)
smoothing

π.χ. $J(f) = \int f''(x) dx$

$J(f) = \int f^2(x) dx = \|f\|_2^2$



$$\min_x \{ R(x) : g(x) \leq c \} \Leftrightarrow$$

$$\min_{x, \mu} L(x, \mu) = R(x) - \mu(g(x) - c)$$

μ : Lagrange multiplier

Το αρχικό πρόβλημα μπορεί να εκφραστεί ως

$$\min \{ \text{RSS}(f) + \underbrace{\lambda J(f)}_{\text{penalty}} \}$$

λ : penalty parameter.

Ridge, Lasso, Elastic net methods

② Basis functions (Συνθετικές λωρύ)

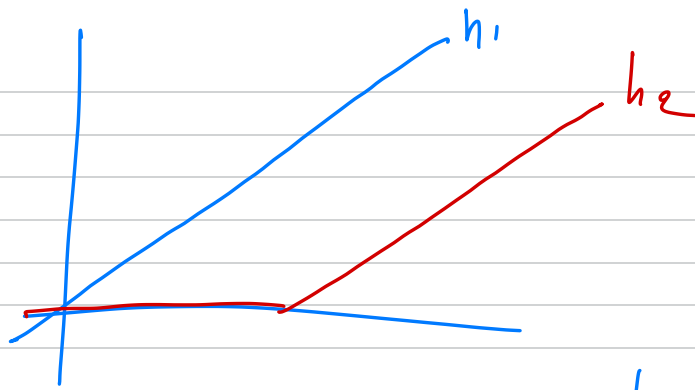
$$\textcircled{1} \quad \hat{f}(x) = \sum_{m=1}^M \theta_m h_m(x)$$

$$\begin{cases} h_j(x), j=1, \dots, M : \text{basis functions} \\ \theta_1, \theta_2, \dots, \theta_M : \text{παράμετροι} \end{cases}$$

Splines

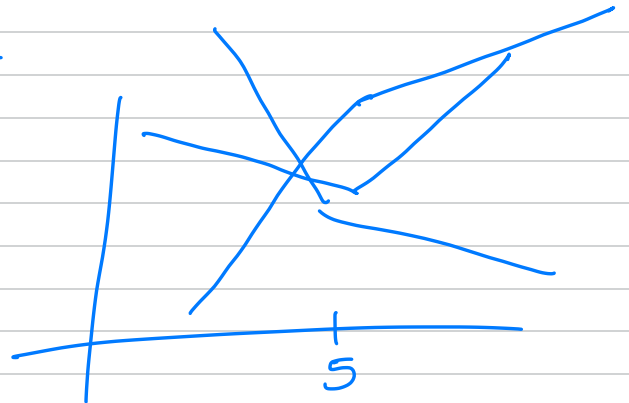
πχ. $h_1(x) = x$

$$h_2(x) = (x-5)^+ = \max(x-5, 0)$$

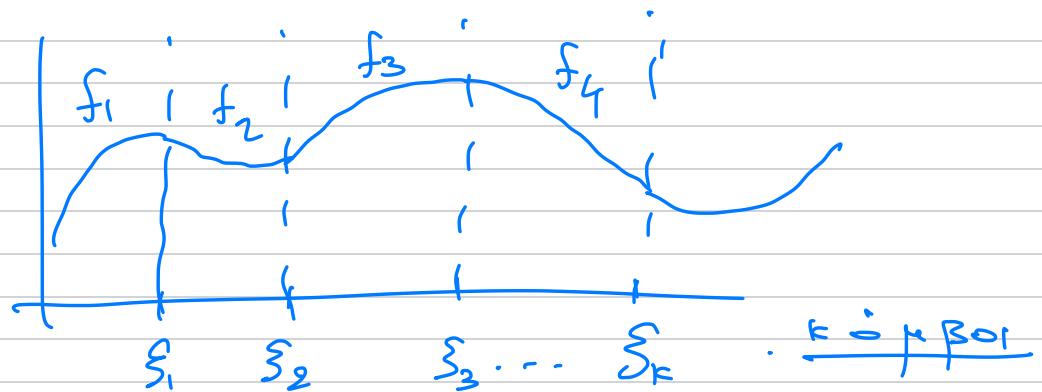


$$h = \theta_1 h_1 + \theta_2 h_2 \Rightarrow ?$$

$$h = \begin{cases} \theta_1 x & , x \leq 5 \\ (\theta_1 + \theta_2) x & x > 5 \end{cases}$$



Γενικά

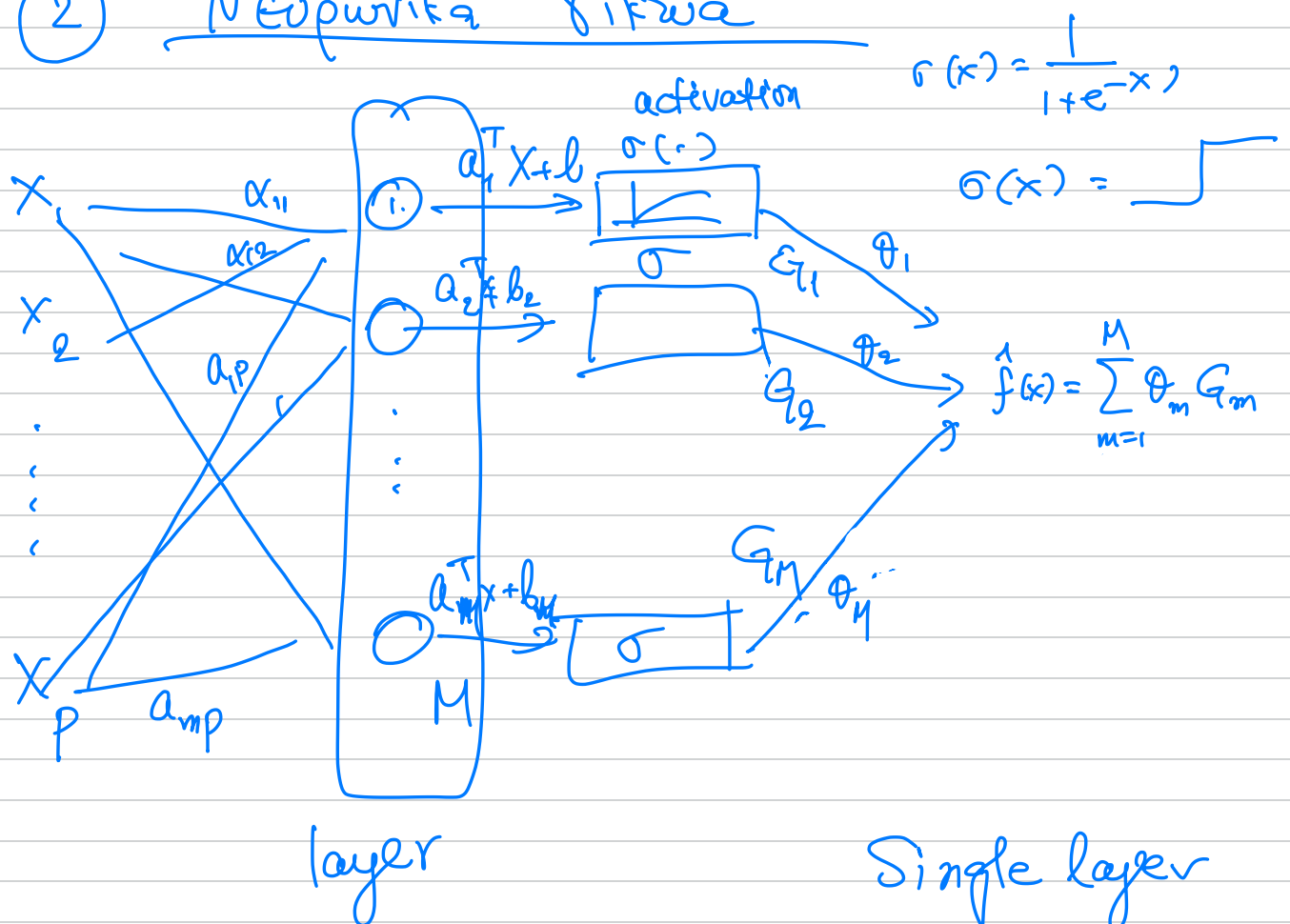


$f_1, f_2, \dots$ υποβλήματα (π.χ. 3^ο βαθμίου)

$\left. \begin{array}{l} \kappa' \quad f() \text{ συνεχής} \\ f'() \quad " \\ f''() \quad " \end{array} \right\} \begin{array}{l} \text{παθρών} \\ \text{περιορισμοί ομαλότητας} \end{array}$

$$f = \begin{cases} f_1 & : x \in (0, \xi_1) \\ \vdots & \\ f_k & : x \in (\xi_{k-1}, \xi_k) \end{cases}$$

2) Νευρωνική Στρώμα



$$\hat{f}(x) = \sum_{m=1}^M \theta_m \sigma(a_m^T x + b_m)$$

basis functions

παραμέτρους $\begin{cases} \theta_1, \dots, \theta_M \in \mathbb{R} \\ b_1, \dots, b_M \in \mathbb{R} \\ a_1, \dots, a_M \in \mathbb{R}^p \end{cases}$

$$\min \text{RSS}(f) = \sum_{i=1}^n (y_i - f(x_i))^2$$

πχ. $p=10$, $M=5$

$$(p+2)M \text{ παράμετροι} = 60$$