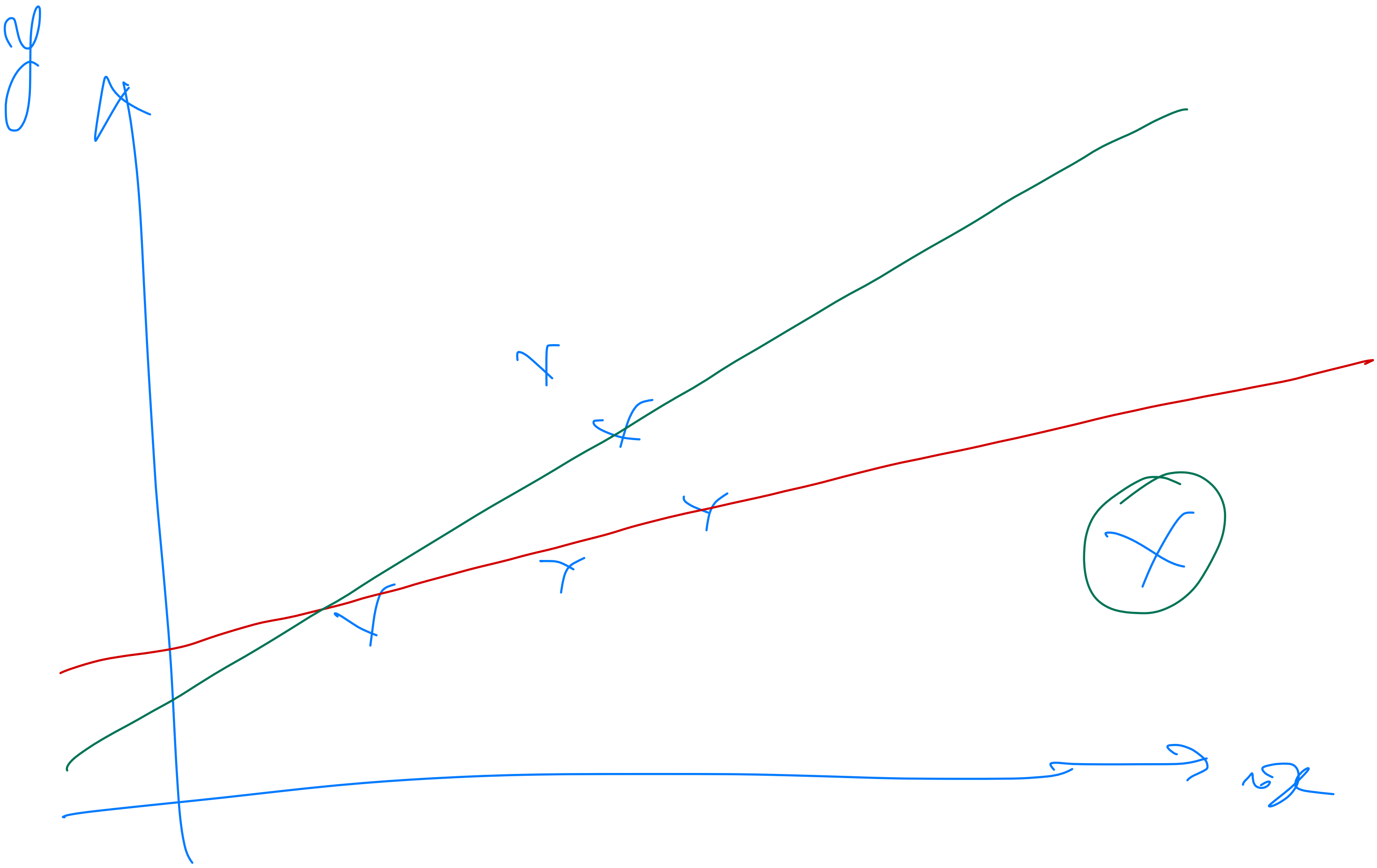
29/9/2021



Παλινδρόμηση

k-NN

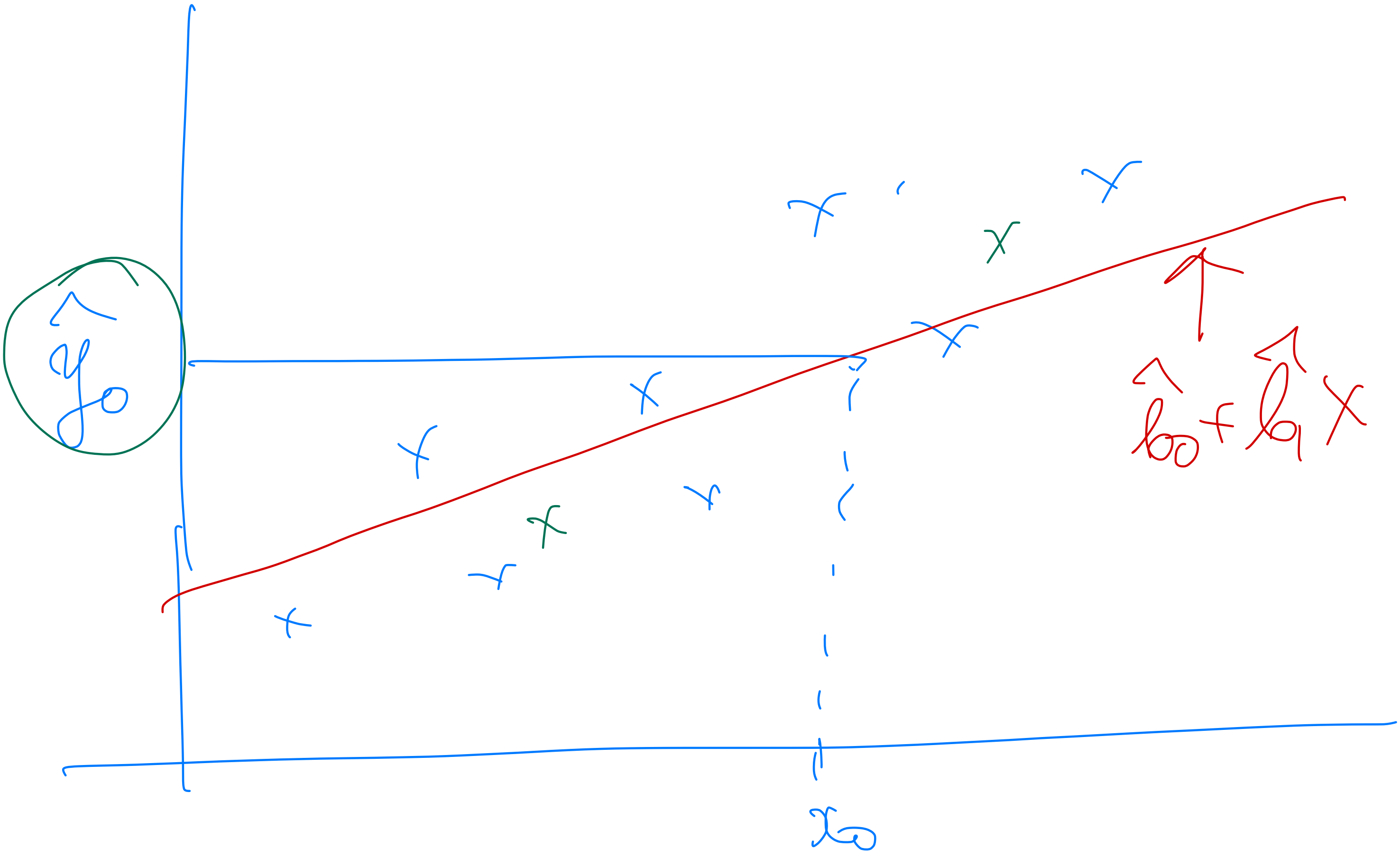
Υπολογιστικά

Υπολογιστικά

Global

Μοντέλα

Local

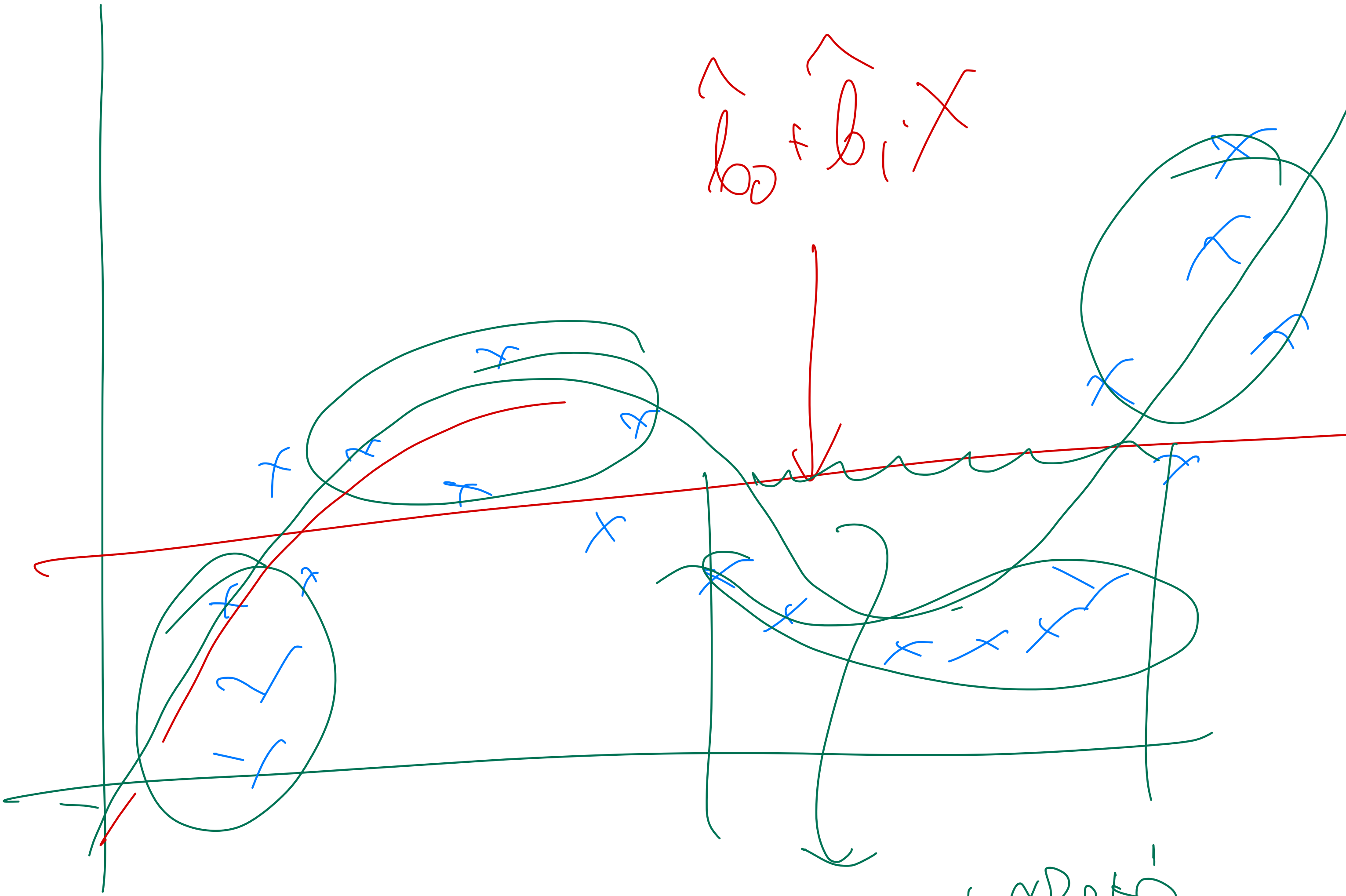


1

$Y = b_0 + b_1 X$

$\hat{y}_0 = \hat{b}_0 + \hat{b}_1 \cdot x_0$

μικρή διασπορά στις προβλέψεις

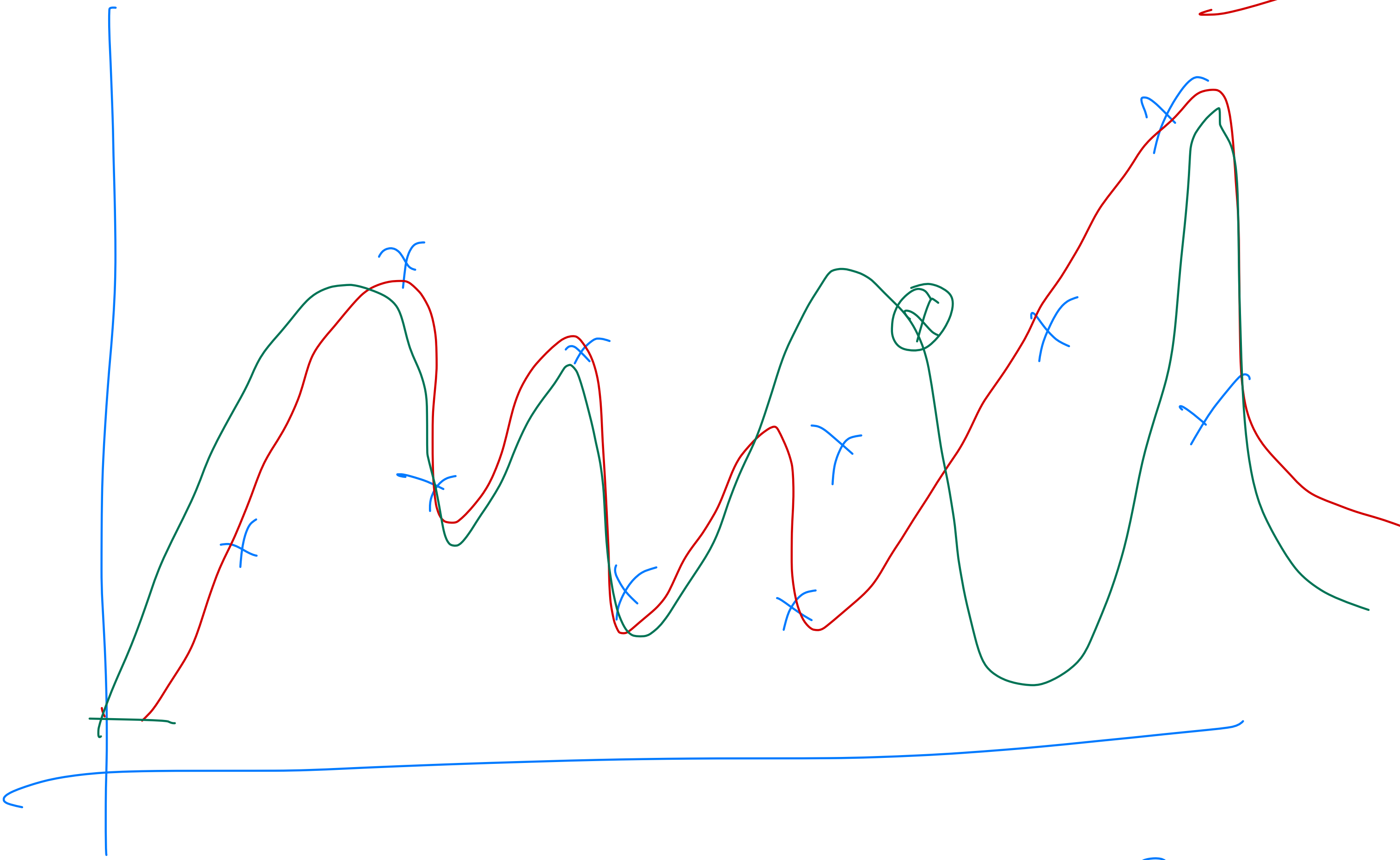


μεγάλο bias
μικρό variance

υποεκπαιδευτό
σφάλμα
(bias)

② $Y = b_0 + b_1 X + b_2 X^2 + b_3 X^3$

$n > 20$

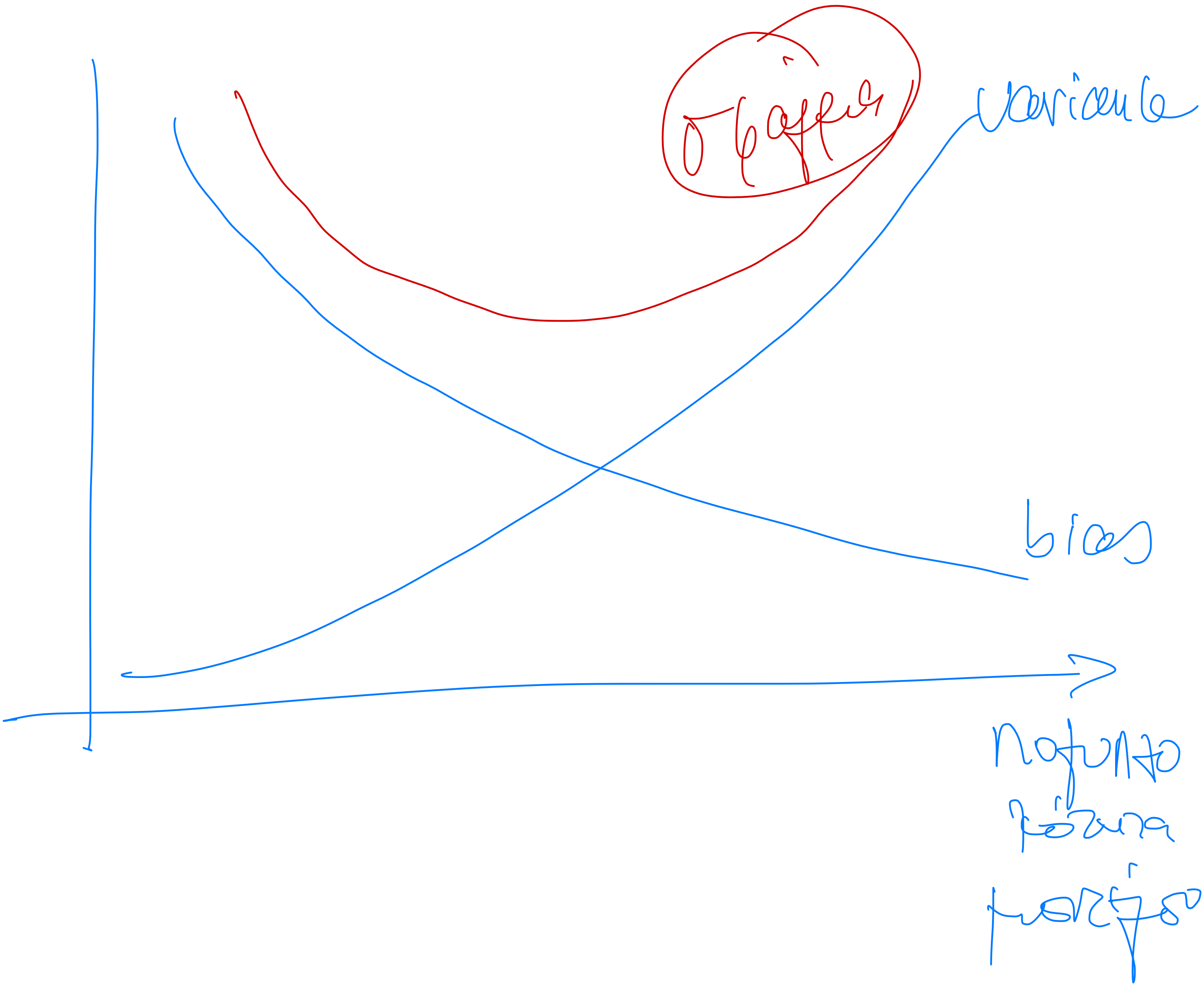


$$Y = b_0 + b_1 X + b_2 X^2 + \cdots + b_{20} X^{20} + \varepsilon$$

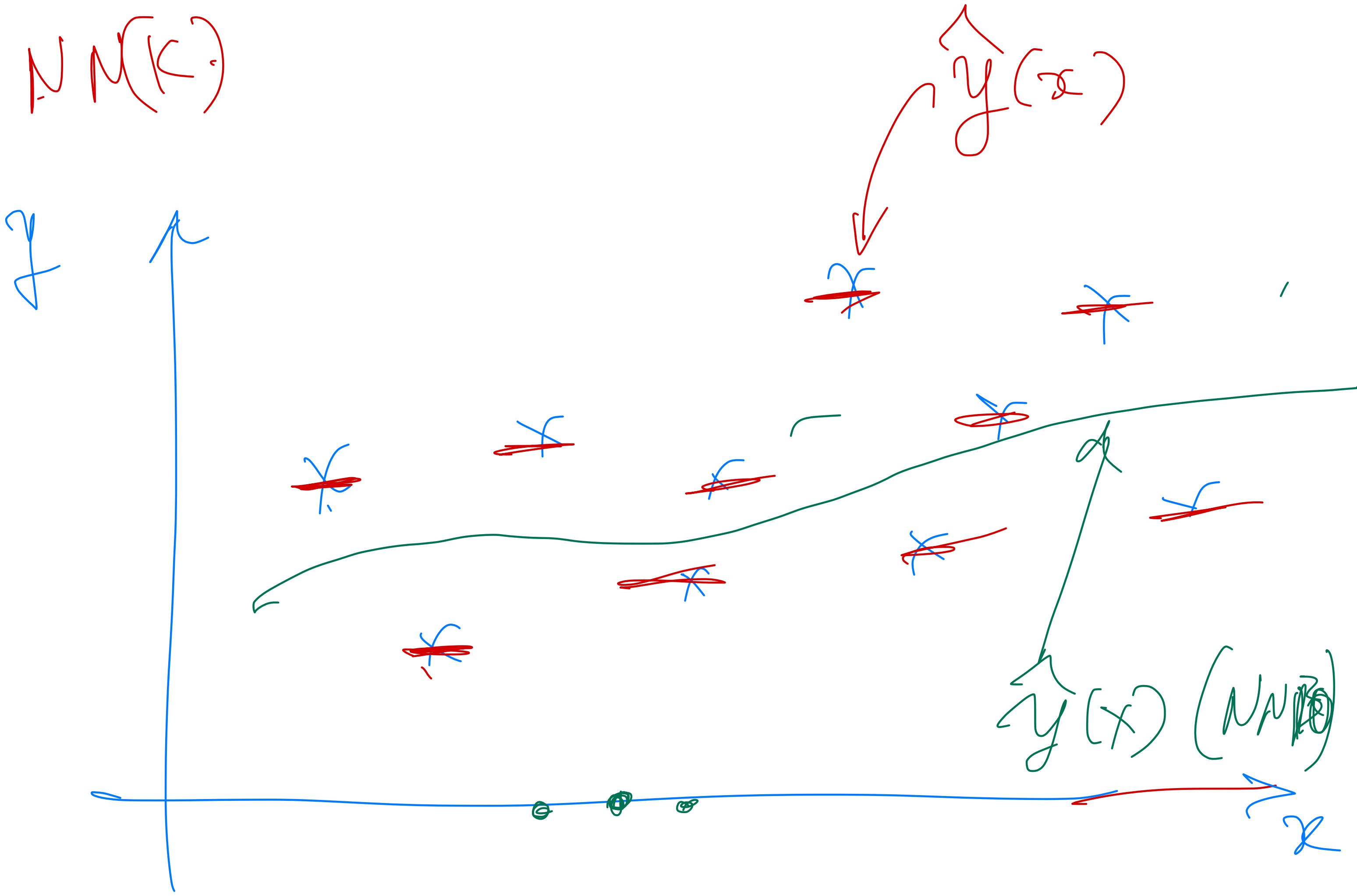
$$\varepsilon \sim N(0, \sigma^2)$$

μικρό: bias

πολύ μεγάλο variance


σφάλμα
variance
bias
πολυπλοκότητα μοντέλου

NN(K)



NN(1) πολύπλοκο μοντέλο

NN(10) απλό μοντέλο

Classification

Εξαρτημένη μεταβλητή κατηγορική : G

$$G \in \{g_1, g_2, \dots, g_K\}$$

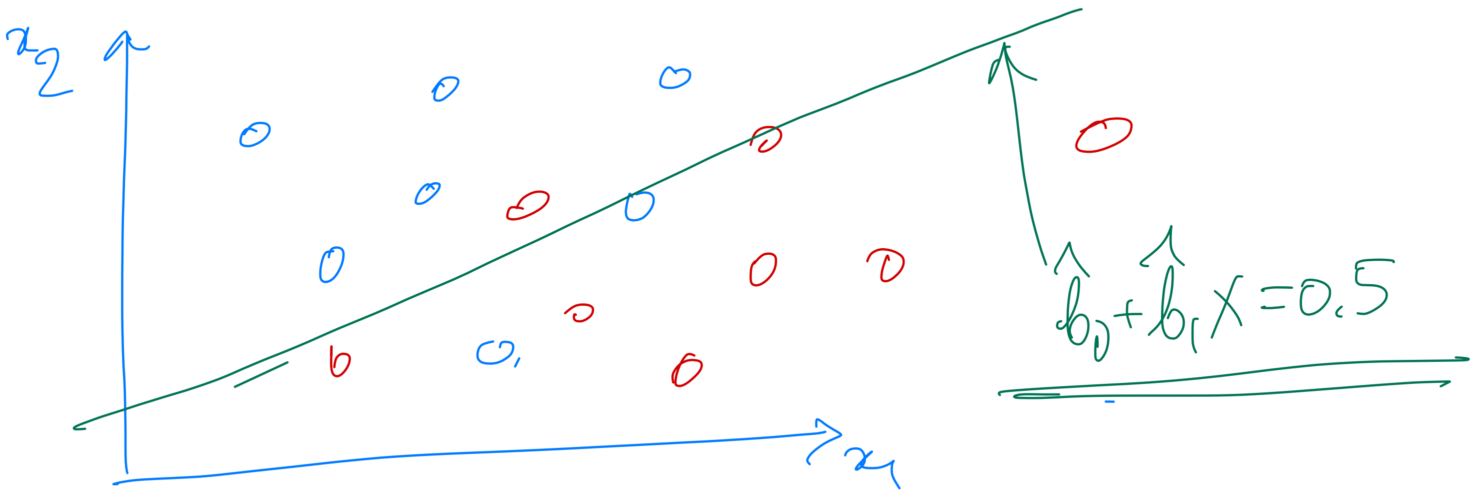
group/class/...

Περίπτωση : $G \in \{0, 1\}$

Logistic regression

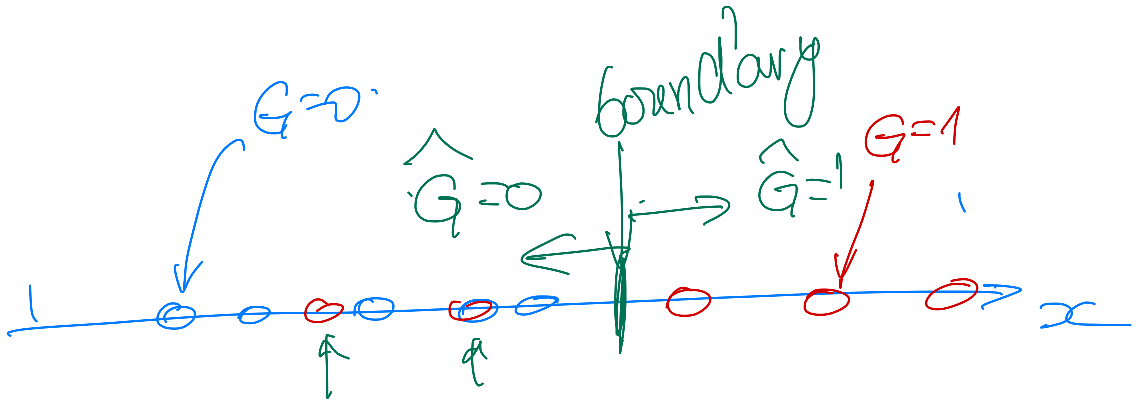
$G \sim \text{Bernoulli}(p)$

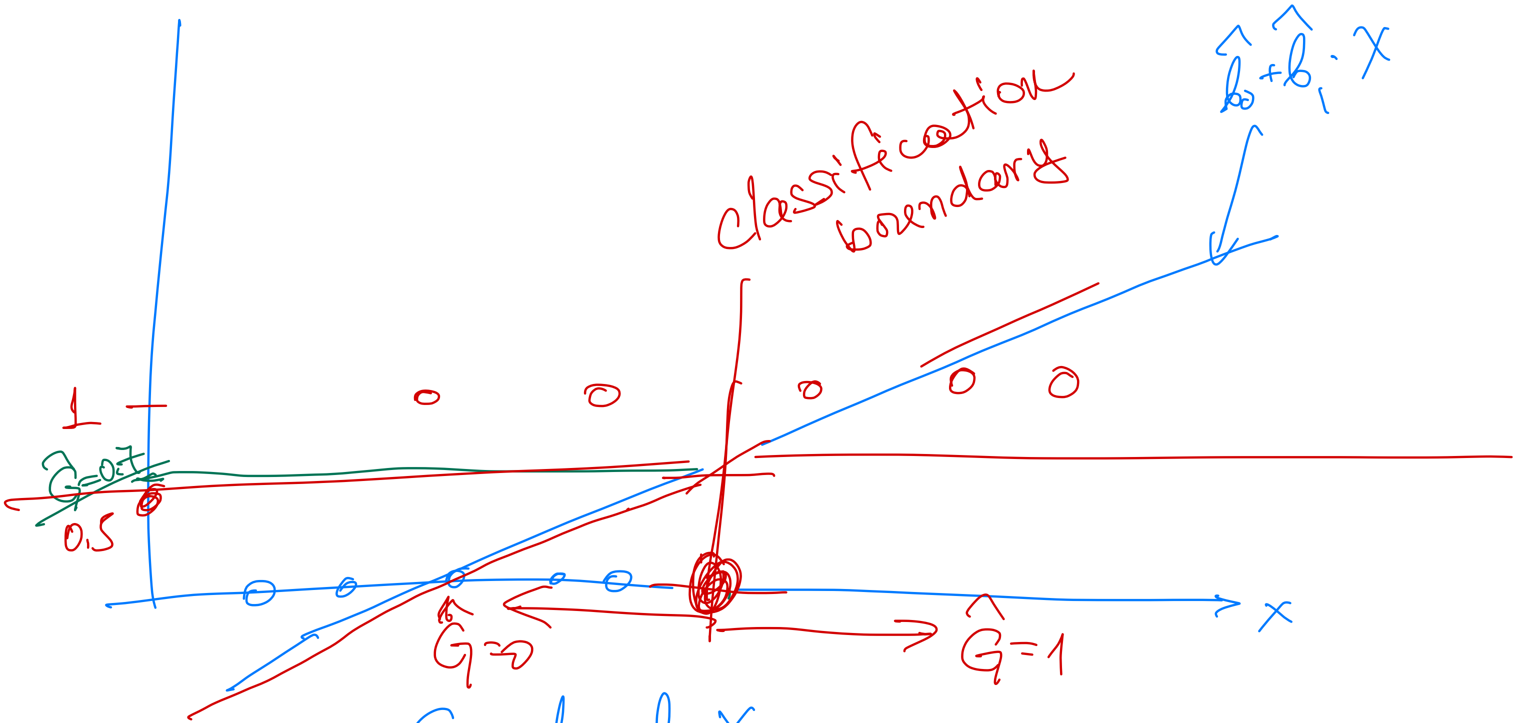
$$\boxed{p = p(\underline{x}) = \frac{e^{\beta^T x}}{1 + e^{\beta^T x}}}$$



$$G = \begin{cases} 0 & \text{μπλέ} \\ 1, & \text{κόκκινο} \end{cases}$$

$$\hat{G} = \underline{b_0 + b_1 \cdot X} + b_2 X^2$$





$$G = b_0 + b_1 X$$

a_0 $\quad \hat{b}_0 + \hat{b}_1 x_0 < 0.5 \Rightarrow \hat{G}_0 = 0$

$\qquad\qquad\qquad\quad > 0.5 \Rightarrow \hat{G}_0 = 1$

Στατιστική Θεωρία Αποφάσεων

predictors/features (ανεξ. μεταβλητές) $X \in \mathbb{R}^p$

διαν. τυχαία μεταβ.

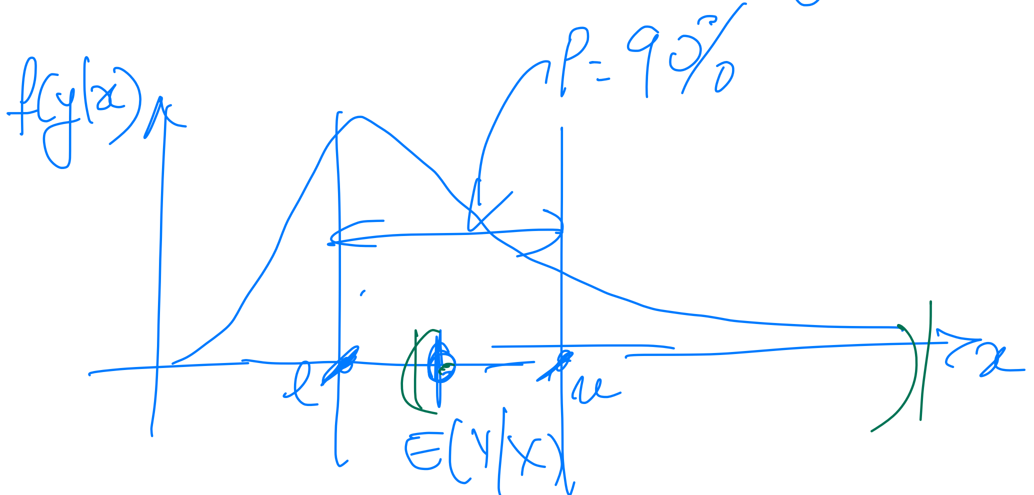
$Y \in \mathbb{R}$ τ.μ. outcome/εξαρτημένη (response)

$P(X, Y)$ joint distribution

$X \longrightarrow f(X)$: predictor του Y

Μια απλή πρόβλεψη

Given $X = x \Rightarrow Y \mid X = x$ $\quad F(y|x)$



Συνάρτηση απώλειας

$L(Y, f(x))$: απώλεια (ποινή) αν η πραγμ. τιμή είναι Y κ' η πρόβλεψη $f(x)$

$L \geq 0 \qquad L = 0 \iff Y = f(X)$

π.χ. $\quad L = (Y - f(x))^2 \leftarrow$

$L = |Y - f(x)| \quad \ldots$

Me $\quad L = (Y - f(X))^2$

$$\text{EPE} = E_P\left(L(Y, f(X))\right)$$

exp. pred. error

$$= E\,(Y - f(X))^2$$

$$\text{EPE}(f) = \int (y - f(x))^2 \, dP(x, y)$$

$$\uparrow$$

"min f"

$$E(Y - f(X))^2 = E_X \boxed{E_{Y|X}\left[(Y - f(X))^2\right]}$$

$$= \int_x \boxed{E_{Y|X=x}\left(Y - f(x)\right)^2} \, dF_X(x)$$

$$\downarrow$$

$\min\limits_{f(x)}$

Αν $\forall x \in \mathbb{R}^p$

$$\min_c E_{Y|X=x}(Y-c)^2 \Rightarrow c^*, \quad f(x) = c^*$$

Τότε η $f(X)$ που προκύπτει με αυτό τον τρόπο ελαχιστοποιεί το $EPE(f)$.

$\forall x$: $h(c, x) = E\left((Y-c)^2 \mid X=x\right)$

$$= E\left(Y^2 - 2cY + c^2 \mid X=x\right)$$

$$= E(Y^2 \mid X=x) - 2c\,\underline{E(Y \mid X=x)} + c^2$$

---

min για $\boxed{c^* = E(Y|X=x)}$

$\boxed{f(x) = E(Y|X=x) \text{ regression function}}$

Αν $L = |Y - f(x)|$

$$f(x) = \text{median}(Y|X=x)$$

$$f(x) = E(Y|X=x)$$ άγνωστη συνάρτηση

Statistics
M Learning
Applied Math

function approximation