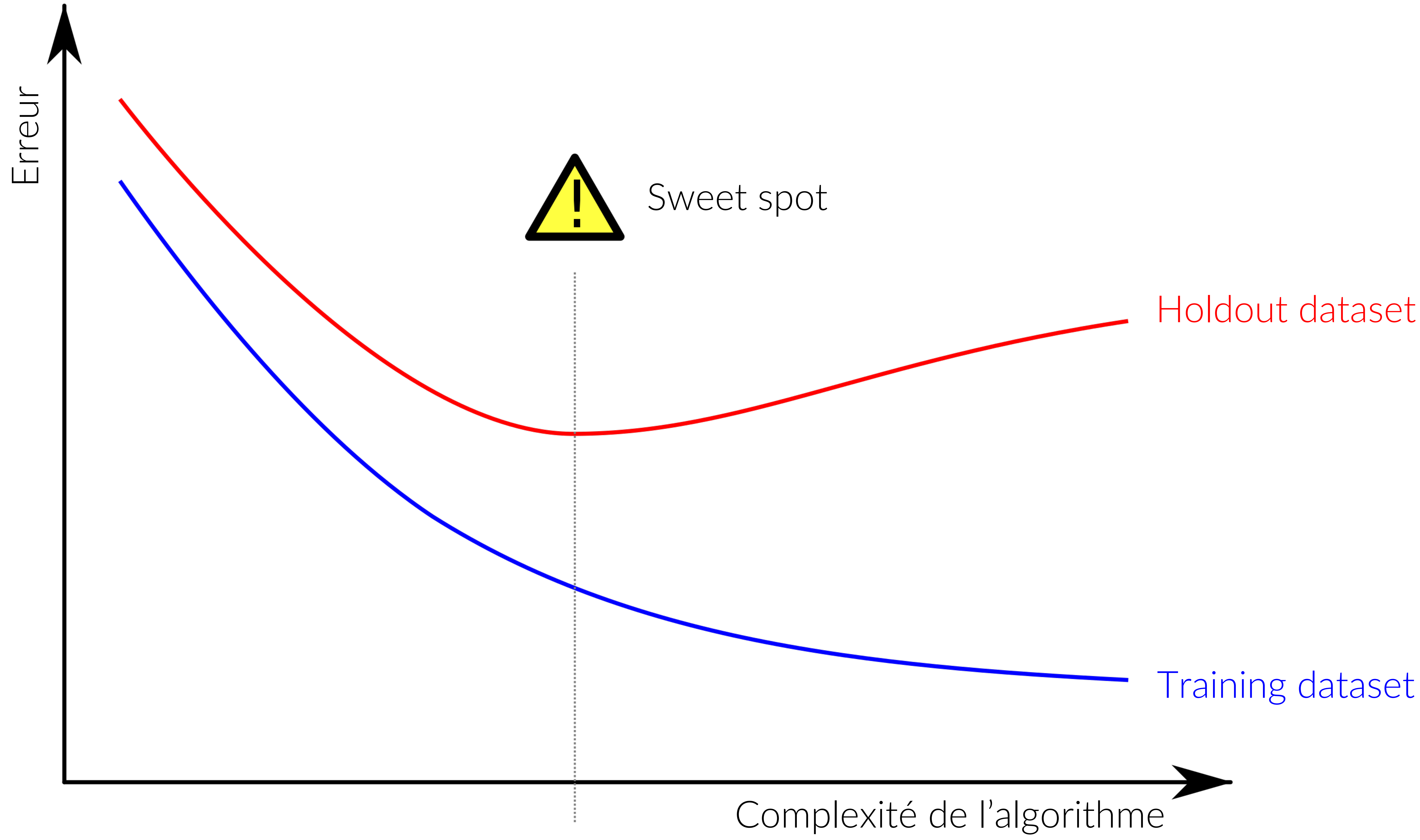
5. Control overfitting

Lev Kiwi

Rappel

Fitting Graph



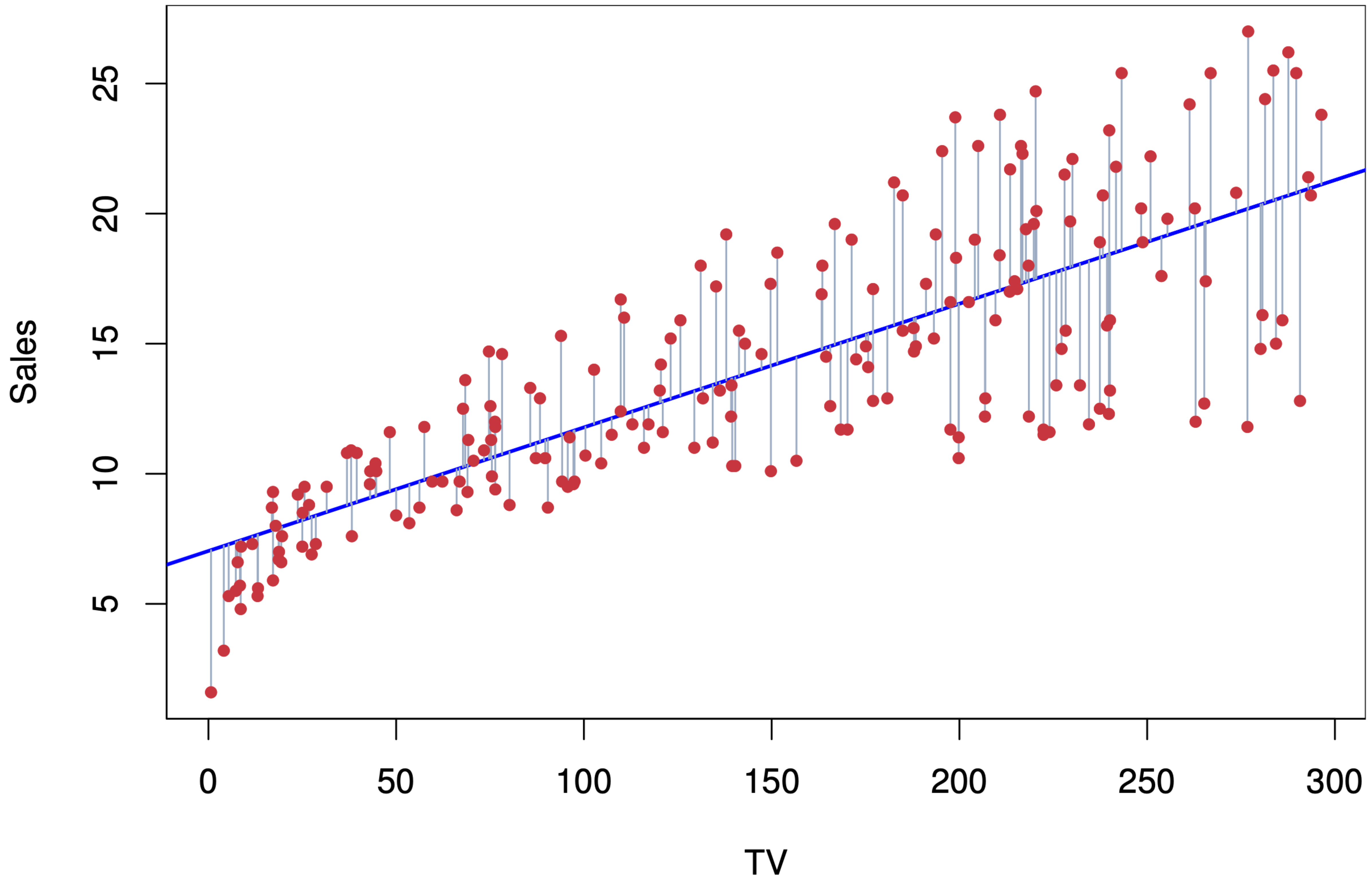
Régression Linéaire

Modèle

L'algorithme cherche les coefficients $a_0, a_1, a_2, \dots, a_n$ telle qu'on minimise l'erreur quadratique sur le dataset d'entrainement,

$y = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n + \varepsilon$ (hyperplan)

$\varepsilon = \sqrt{\sum_i \varepsilon_i{}^2}$ (erreur quadratique)



Lasso Regression

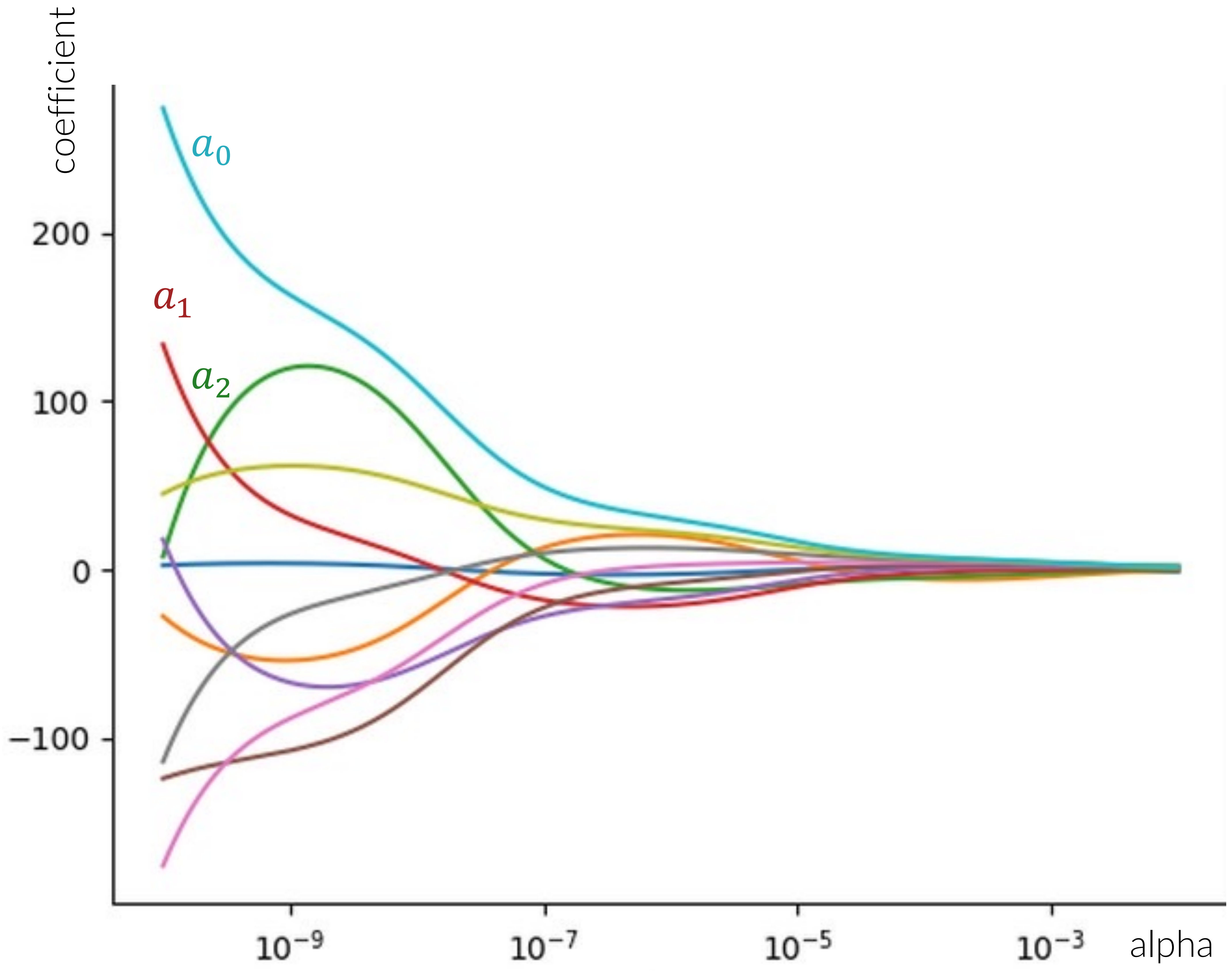
Modèle

L'algorithme cherche les coefficients $a_0, a_1, a_2, \dots, a_n$ telle qu'on minimise la fonction objectif F sur le dataset d'entrainement,

$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \varepsilon$ (hyperplan)

$\mathrm{F} = \sqrt{\sum_i {\varepsilon_i}^2} + \alpha \sum_k |a_k|$ (objective function)

Le coefficient α est un coefficient de force de la régularisation. Lorsque $\alpha = 0$ on retrouve la régression linéaire classique.



Lasso, Ridge Regression & elastic Net

Modèle

L'algorithme cherche les coefficients $a_0, a_1, a_2, \dots, a_n$ telle qu'on minimise la fonction objectif F sur le dataset d'entrainement,

$$y = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \varepsilon \qquad \text{(hyperplan)}$$

Les modèles ont toutes des Loss function différentes dans leurs fonctions objectifs.

Les valeurs α, λ_1, λ_2 sont des hyperparamètres de ces modèles.

Lasso Regression

$$\mathrm{F} = \sqrt{\textstyle\sum_i {\varepsilon_i}^2} + \alpha \textstyle\sum_k |a_k| \qquad \text{(L1 Loss)}$$

Ridge Regression

$$\mathrm{F} = \sqrt{\textstyle\sum_i {\varepsilon_i}^2} + \alpha \sqrt{\textstyle\sum_k {a_k}^2} \qquad \text{(L2 Loss)}$$

Elastic Net

$$\mathrm{F} = \sqrt{\textstyle\sum_i {\varepsilon_i}^2} + \lambda_1 \textstyle\sum_k |a_k| + \lambda_2 \sqrt{\textstyle\sum_k {a_k}^2} \qquad \text{(L1 + L2 Loss)}$$