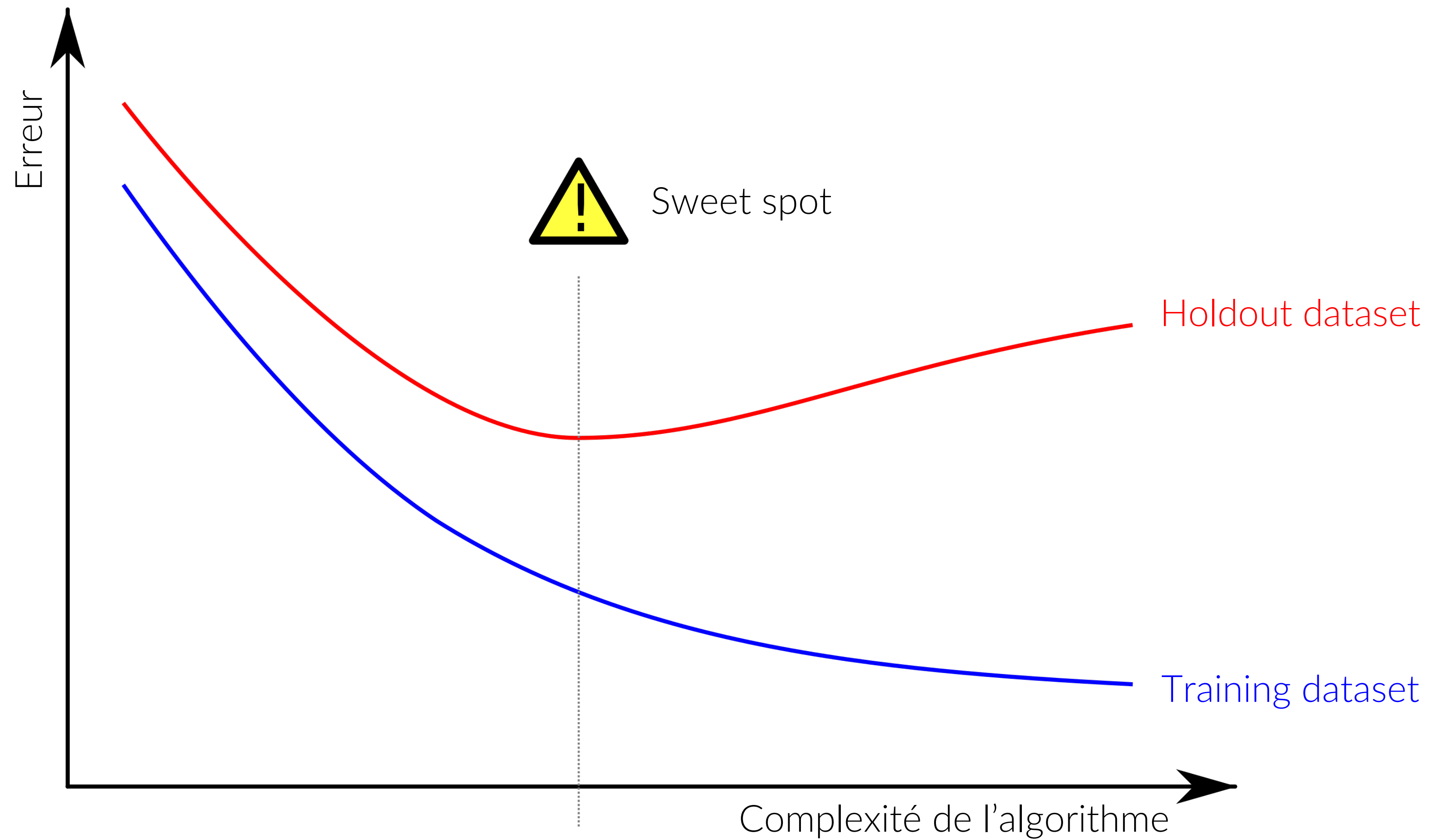


5. CONTROL OVERFITTING

LEV KIWI

FITTING GRAPH



RÉGRESSION LINÉAIRE

Modèle

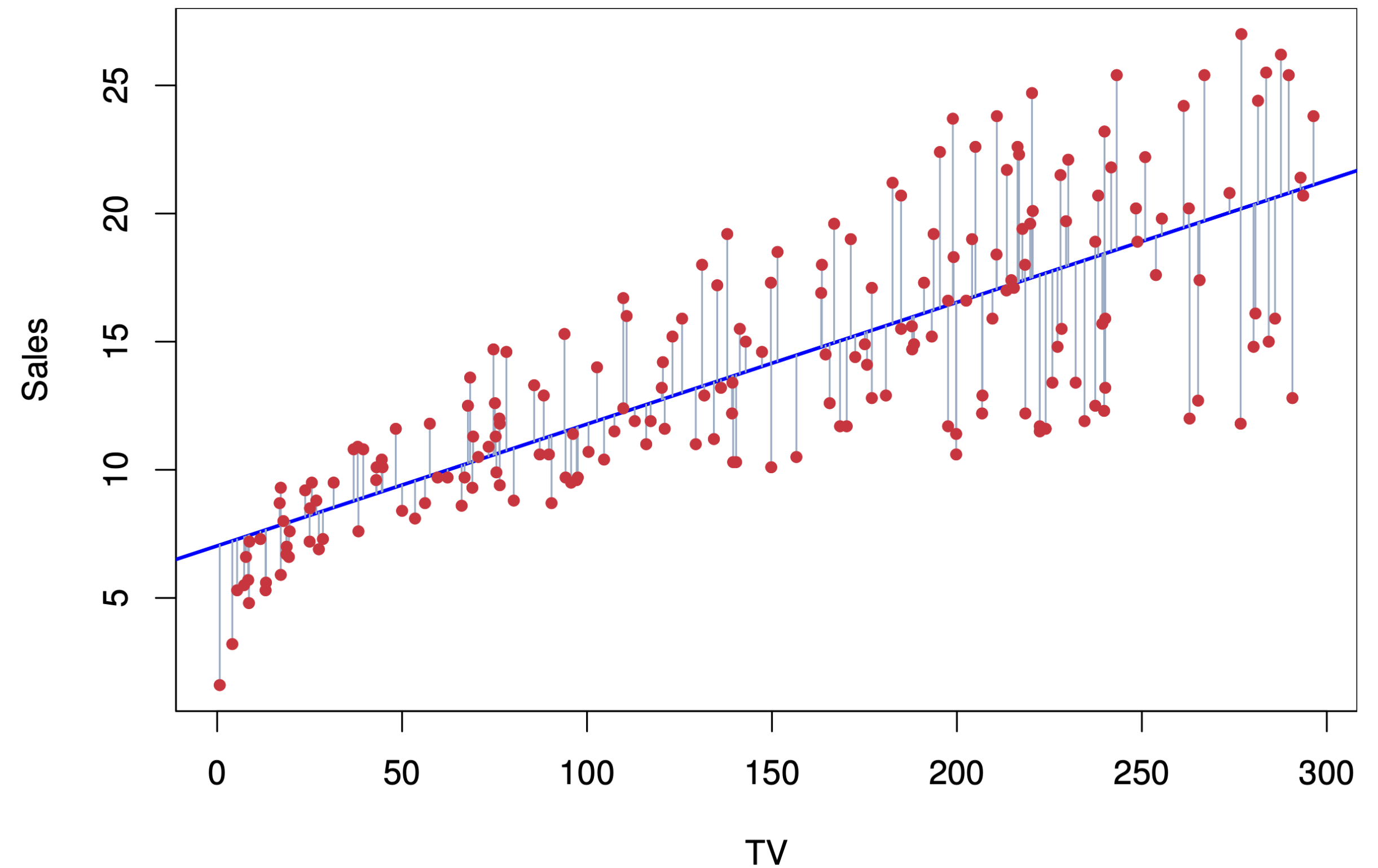
L'algorithme cherche les coefficients $a_0, a_1, a_2, \dots, a_n$ telle qu'on minimise l'erreur quadratique sur le dataset d'entraînement,

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + \varepsilon$$

(hyperplan)

$$\varepsilon = \sqrt{\sum_i \varepsilon_i^2}$$

(erreur quadratique)



LASSO REGRESSION

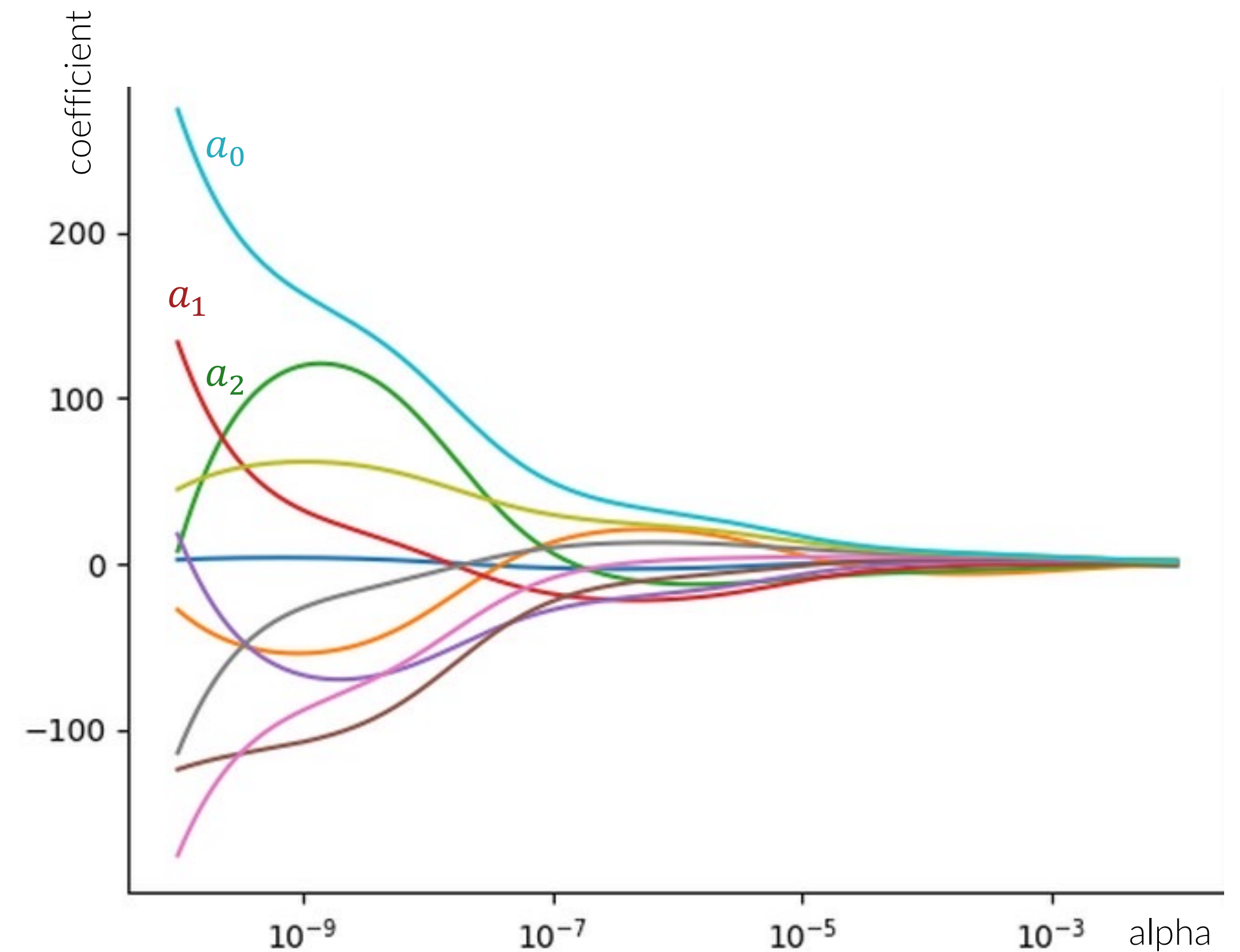
Modèle

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'entraînement,

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + \varepsilon \quad (\text{hyperplan})$$

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

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'entraînement,

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + \varepsilon \quad (\text{hyperplan})$$

Les modèles ont toutes des **fonctions de régularisation** différentes dans leurs fonctions objectifs.

Les valeurs $\alpha, \lambda_1, \lambda_2$ sont des **hyperparamètres** de ces modèles.

Lasso Regression

$$F = \sqrt{\sum_i \varepsilon_i^2} + \alpha \sum_k |a_k| \quad (\text{L1 Regularization})$$

Ridge Regression

$$F = \sqrt{\sum_i \varepsilon_i^2} + \alpha \sqrt{\sum_k a_k^2} \quad (\text{L2 Regularization})$$

Elastic Net

$$F = \sqrt{\sum_i \varepsilon_i^2} + \lambda_1 \sum_k |a_k| + \lambda_2 \sqrt{\sum_k a_k^2} \quad (\text{L1 + L2 Reg.})$$