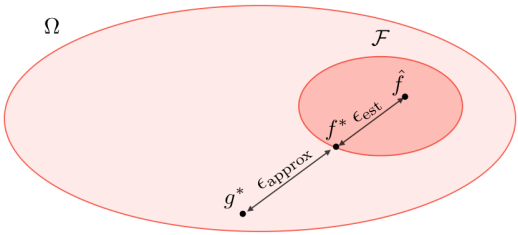
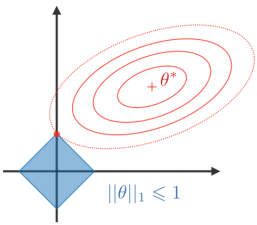
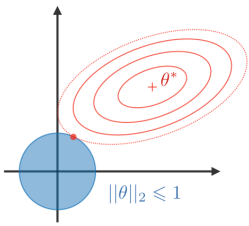
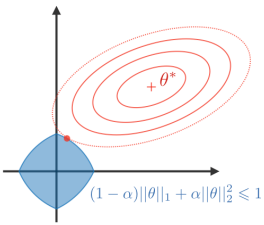
d'estimation ϵ_{est} quantifie la qualité du prédicteur $\hat{f}$ par rapport au meilleur prédicteur f^* de la classe d'hypothèses $\mathcal{F}$.



❒ **Régularisation** – Le but de la régularisation est d'empêcher le modèle de surapprendre (en anglais *overfit*) les données en s'occupant ainsi des problèmes de variance élevée. La table suivante résume les différents types de régularisation couramment utilisés :

LASSO	**Ridge**	**Elastic Net**
- Réduit les coefficients à 0 - Bénéfique pour la sélection de variables	Rapetissit les coefficients	Compromis entre sélection de variables et coefficients de faible magnitude
		
$\ldots + \lambda\|\|\theta\|\|_1$ $\lambda \in \mathbb{R}$	$\ldots + \lambda\|\|\theta\|\|_2^2$ $\lambda \in \mathbb{R}$	$\ldots + \lambda\left[(1-\alpha)\|\|\theta\|\|_1 + \alpha\|\|\theta\|\|_2^2\right]$ $\lambda \in \mathbb{R}, \quad \alpha \in [0,1]$

❒ **Hyperparamètres** – Les hyperparamètres sont les paramètres de l'algorithme d'apprentissage et incluent parmi d'autres le type de caractéristiques utilisé ainsi que le paramètre de régularisation λ, le nombre d'itérations T le taux d'apprentissage η.

❒ **Vocabulaire** – Lors de la sélection d'un modèle, on divise les données en 3 différentes parties :

Données d'entraînement	**Données de validation**	**Données de test**
- Le modèle y est entrainé - Constitue normalement 80 du jeu de données	- Le modèle y est évalué - Constitue normalement 20 du jeu de données - Aussi appelé données de développement (en anglais *hold-out* ou *development set*)	- Le modèle y donne ses prédictions - Données jamais observées