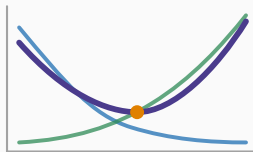


Apprentissage statistique

Chapitre 2 : Le compromis biais-variance



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

La décomposition

Le compromis, chiffré

Ce qui déplace l'optimum

La malédiction de la dimension

Ce qu'il faut retenir

La décomposition

Trois termes, et un seul est incompressible

L'erreur quadratique moyenne, décomposée

$$\text{EQM} = \underbrace{\left[\text{biais}(\hat{f}) \right]^2}_{\text{erreur systématique}} + \underbrace{\text{Var}(\hat{f})}_{\text{instabilité}} + \underbrace{\sigma^2}_{\text{bruit}}$$

Ce que chaque terme signifie

Le **biais** est l'erreur commise en approchant un phénomène complexe par un modèle trop simple. Il mesure une **rigidité**.

La **variance** est la sensibilité du modèle à l'échantillon : un autre tirage donnerait-il un modèle très différent ? Elle mesure une **instabilité**.

Le **bruit** est irréductible.

L'image du tir

Un tireur **biaisé** groupe ses impacts, mais loin du centre. Un tireur à forte **variance** vise juste en moyenne et disperse.

Le premier a besoin d'un modèle plus flexible ; le second, de plus de contrainte. **Ce sont deux maladies opposées, et le remède de l'une aggrave l'autre.**

Le compromis, chiffré

Un modèle de risque de crédit, à complexité croissante

Complexité	Biais ²	Variance	Bruit	EQM
1	9,0	0,5	1	10,5
3	4,0	1,0	1	6,0
5	1,0	2,0	1	4,0
10	0,5	5,0	1	6,5
20	0,2	12,0	1	13,2

Le biais **décroît** sans cesse, la variance **croît** sans cesse. Leur somme passe par un minimum, ici en complexité 5, pour une EQM de 4.

Définition

	Biais	Variance	Erreur d'apprentissage	Erreur de test
Sous-ajustement	élevé	faible	élevée	élevée
Sur-ajustement	faible	élevée	très faible	élevée

Intuition

Une erreur d'apprentissage **élevée** signale un sous-ajustement : le modèle ne capte même pas ce qu'il a sous les yeux.

Une erreur d'apprentissage **quasi nulle** avec une erreur de test élevée signale un sur-ajustement : le modèle a mémorisé le bruit.

C'est l'écart entre les deux erreurs qui diagnostique, non leur niveau.

Ce qui déplace l'optimum

La taille de l'échantillon

Plus n est grand, plus la variance diminue à complexité donnée. L'optimum se déplace donc vers **plus de complexité**.

Un modèle très flexible n'est pas dangereux en soi : il l'est **relativement au nombre d'observations**.

Le nombre de variables

Ajouter des variables augmente la flexibilité, donc la variance. En grande dimension, ce terme devient dominant — c'est le sujet du chapitre 5.

La régularisation

On peut **brider** un modèle flexible plutôt que d'en choisir un simple : accepter un peu de biais pour retrancher beaucoup de variance.

C'est le principe de ridge et du lasso, au chapitre 4, et de presque toutes les méthodes modernes.

Application en gestion

Le chapitre 17 d'*Économétrie avancée* traite de la **variable omise** : omettre une variable pertinente **biaise** les coefficients.

En prévision, ce même arbitrage change de sens. Omettre une variable faiblement utile **augmente le biais** mais **réduit la variance**, et peut améliorer la prévision.

L'économétrie cherche un estimateur sans biais. **L'apprentissage statistique accepte du biais s'il achète de la variance en échange**. C'est le renversement central de ce cours.

La malédiction de la dimension

Pourquoi l'intuition géométrique trompe

Malédiction de la dimension

Quand le nombre de variables augmente, le volume de l'espace croît exponentiellement, et les observations s'y **isolent**.

À dimension élevée, tous les points deviennent à peu près équidistants les uns des autres.

Un ordre de grandeur

Pour couvrir 10 % du volume d'un cube unité autour d'un point, il faut un côté de :

Dimension	Côté nécessaire
1	0,10
2	0,32
10	0,79

En dimension 10, un « voisinage » de 10 % du volume couvre déjà près de 80 % de l'étendue de chaque variable. Il n'a plus rien de local.

Ce qu'il faut retenir

Cinq points

1. $EQM = \text{biais}^2 + \text{variance} + \text{bruit}$.
2. La flexibilité **réduit le biais** et **augmente la variance**; l'optimum est intérieur — ici complexité 5, EQM 4.
3. **Sous-ajustement** : les deux erreurs élevées. **Sur-ajustement** : apprentissage quasi nul, test élevé.
4. Plus n est grand, plus on peut se permettre de complexité.
5. En prévision, **accepter du biais pour retrancher de la variance** est légitime — contrairement à l'économétrie.

La suite

Le compromis se règle sur l'erreur de test, qu'on ne connaît pas. Le chapitre 3 donne le moyen de l'estimer sans y toucher.