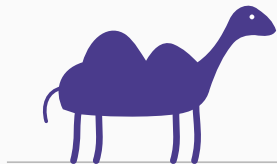


Introduction à l'analyse des données

Chapitre 7 : L'inertie, ou l'information d'un nuage



Jaouad Madkour

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

Mesurer l'information

Le résultat qui relie tout

Décomposer l'inertie

Ce qu'il faut retenir

Mesurer l'information

La question laissée en suspens au chapitre 1

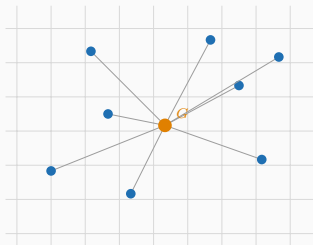
Le photographe cherche l'angle qui **perd le moins d'information**. Encore faut-il savoir la mesurer : sans quoi on ne peut ni comparer deux angles, ni dire qu'une représentation est bonne.

L'idée

Un nuage dont tous les points seraient **confondus** ne contiendrait aucune information : tous les individus seraient identiques.

Un nuage **très dispersé** en contient beaucoup : les individus diffèrent, et ces différences sont précisément ce qu'on veut décrire.

L'information d'un nuage, c'est donc sa dispersion.



L'inertie totale

$$I_{\text{tot}} = \frac{1}{n} \sum_{i=1}^n d^2(i, G)$$

Elle mesure exactement ce que mesurait la variance — et pour une seule variable, elle *est* la variance.

Sur données centrées-réduites, chaque variable apporte une inertie de 1. Avec p variables, l'inertie totale vaut donc p : ici 8.

C'est la monnaie de tout le cours. « Cet axe explique 56,7 % » signifiera toujours : il en conserve 56,7 %.

Inertie totale

$$I = \frac{1}{n} \sum_{i=1}^n d^2(i, G)$$

C'est la **moyenne des carrés des distances** des individus au centre de gravité.

Le mot vient de la mécanique

En physique, le moment d'inertie mesure la difficulté à faire tourner un corps autour d'un axe : plus la masse est éloignée de l'axe, plus l'inertie est grande.

Ici, plus les individus sont éloignés du centre, plus l'inertie est grande. **Le vocabulaire est emprunté, la notion est la même.**

Le résultat qui relie tout

L'inertie est une somme de variances

Le théorème

$$I = \sum_{j=1}^p V(x_j)$$

L'inertie totale du nuage est la somme des variances des variables.

Pourquoi c'est immédiat

La distance au carré d'un individu au centre est, par Pythagore, la somme sur les p variables des écarts au carré à chaque moyenne.

Moyenner sur les individus donne, variable par variable, exactement la variance. **L'inertie est donc la dispersion totale, tous critères confondus** — la dispersion du chapitre 3 est la taille du nuage du chapitre 6.

Exemple

$$I = V(x) + V(y) = 8 + 32 = 40$$

La fidélité pèse quatre fois plus que l'accueil dans cette inertie — uniquement parce qu'elle est notée avec plus de dispersion.

Le cas des données réduites

Le résultat décisif

Après réduction, chaque variable a une variance de 1. Donc

$$I = p$$

L'inertie totale d'un tableau réduit est égale au nombre de variables.

Sur le fil rouge

Deux variables réduites : $I = 1 + 1 = 2$.

L'accueil et la fidélité pèsent désormais **autant l'une que l'autre**, ce qui était bien l'intention du chapitre 3.

Pourquoi ce résultat est si commode

Il donne un **étalon** : si l'inertie totale vaut p , dire qu'un axe en porte 1,9 sur 2 signifie immédiatement 95 %.

Sans lui, « 1,9 » ne voudrait rien dire. C'est ce qui rendra lisibles les pourcentages du chapitre 10.

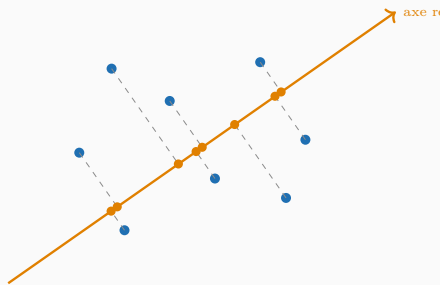
Décomposer l'inertie

Ce qu'une projection conserve

Le théorème de Pythagore, appliqué au nuage

Projeter le nuage sur un axe décompose l'inertie en deux parts :

$$\underbrace{I}_{\text{totale}} = \underbrace{I_{\text{projetée}}}_{\text{conservée}} + \underbrace{I_{\text{résiduelle}}}_{\text{perdue}}$$



Projeter, c'est écraser.

Chaque point bleu tombe sur l'axe. Ce qui subsiste : sa position *le long* de l'axe. Ce qui disparaît : sa distance à l'axe — les traits pointillés.

Théorème de Pythagore, appliqué au nuage entier :

$$I_{\text{tot}} = \underbrace{I_{\text{projetée}}}_{\text{conservée}} + \underbrace{I_{\text{résiduelle}}}_{\text{perdue}}$$

Le meilleur axe est donc celui qui maximise l'inertie projetée — ce qui revient exactement à minimiser la somme des carrés des pointillés.

Le chameau, enfin chiffré

L'inertie projetée est ce que la photographie **conserve** : l'inertie résiduelle est ce qu'elle **écrase**

Ce qu'il faut retenir

Cinq points

1. L'**information** d'un nuage est sa **dispersion** : un nuage concentré ne dit rien.
2. L'**inertie** est la moyenne des carrés des distances au centre de gravité.
3. Elle est égale à la **somme des variances** — le pont entre géométrie et statistique.
4. Sur données **réduites**, $I = p$: ici $I = 2$. C'est l'étalon qui rend les pourcentages lisibles.
5. Une projection **décompose** l'inertie en une part conservée et une part perdue.

La suite

Le problème est posé, et il est bien posé : maximiser l'inertie projetée. Le chapitre 8 explique comment on le résout.