Analyse Multivariée

On observe un nombre n d’individus à travers p caractères où p > 2. L’objectif est de trouver les méthodes qui permettent de révéler les différences les plus significatives entre les individus.

On va se restreindre dans un premier temps à p caractères quantitatifs. Nous avons à faire à un nuage d’individus (considérés comme des vecteurs à p dimensions).

On essaie de projeter ce nuage d’individus dans un espace de faible dimension en minimisant la perte d’information (les différences entre individus).

Avant de réaliser le traitement, il faut demander à la machine de faire quelques arrangements :

1. On ne change pas le problème en centrant le nuage (transformation isomorphe)
2. On normalise les données en les divisant par l’écart-type de chaque caractère : pour calculer la ‘distance’ on fait la somme des différences au carré de chaque caractère divisées par l’écart-type :
3. On va cherche l’inertie du nuage par rapport à l’origine (résistance du nuage par rapport à l’origine)
4. Soit u un vecteur tel que . On note la droite vectorielle de vecteur directeur N :

Remarque : le IN2 est nécessairement inférieur eu IN 1.