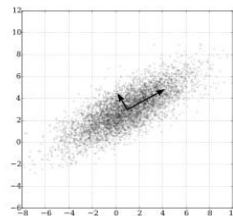


DEVOIR-1 INF2044**Analyse en Composantes Principales (ACP)**

L'**analyse en composantes principales** (ACP ou PCA en anglais pour *principal component analysis*), ou selon le domaine d'application la **transformation de Karhunen–Loève (KLT)**¹, est une méthode de la famille de l'[analyse des données](#) et plus généralement de la [statistique multivariée](#), qui consiste à transformer des variables liées entre elles (dites « corrélées » en statistique) en nouvelles variables décorrélées les unes des autres. Ces nouvelles variables sont nommées « composantes principales », ou axes principaux. Elle permet au praticien de réduire le nombre de variables et de rendre l'information moins redondante.

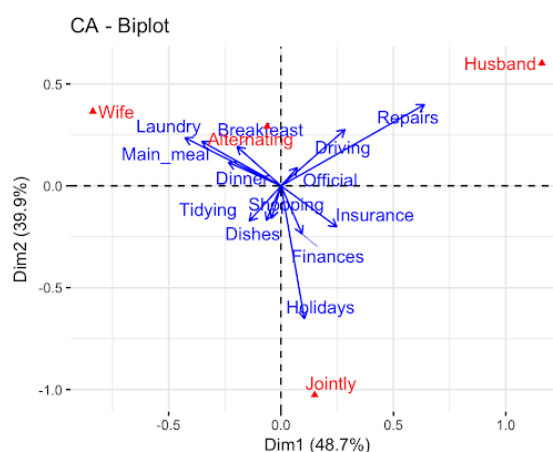
Il s'agit d'une approche à la fois géométrique² (les variables étant représentées dans un nouvel espace, selon des directions d'inertie maximale) et statistique (la recherche portant sur des axes indépendants expliquant au mieux la variabilité — la [variance](#) — des données). Lorsqu'on veut *compresser* un ensemble de variables aléatoires, les *premiers axes* de l'analyse en composantes principales sont un *meilleur choix*, du point de vue de l'inertie ou de la variance.

L'outil mathématique est appliqué dans d'autres domaines que les statistiques et est parfois appelé **décomposition orthogonale aux valeurs propres** ou **POD** ([anglais](#) : *proper orthogonal decomposition*).

**Analyse factorielle des correspondances**

Introduction : Les méthodes d'[analyse factorielle](#) des correspondances (AFC) tout comme celles d'[analyse en composantes principales](#) (ACP) s'utilisent pour décrire et hiérarchiser les relations statistiques qui peuvent exister entre des individus placés en ligne et des variables placées en colonnes dans un tableau rectangulaire de données. L'une et l'autre de ces deux méthodes considèrent le tableau de données comme un nuage de points dans un espace mathématique ayant autant de dimensions qu'il y a de colonnes dans le tableau de données ; elles cherchent à le projeter sur des axes ou des plans (appelés factoriels) de façon que l'on puisse en visualiser et étudier au mieux la forme et donc rechercher globalement des corrélations. La spécificité de l'AFC est qu'elle considère en même temps un nuage de point représentant les lignes (individus) et un autre représentant les colonnes (variables). Les logiciels d'AFC fournissent donc en sortie une ou plusieurs figures de plans factoriels sur lesquels sont positionnés à la fois les individus et les variables.

Principe : Le principe de ces méthodes est de partir sans *a priori* sur les données et de les décrire en analysant la hiérarchisation de l'information présente dans les données. Pour ce faire, les analyses factorielles étudient l'inertie du nuage de points ayant pour coordonnées les valeurs présentes sur les lignes du tableau de données.



DEVOIR-2 INF2044

Problème :

Un état souhaite réduire le taux d'illectronisme dans une zone.

Données disponibles :

Historique des recherches sur le pourcentage de la population n'ayant pas d'accès à internet ou pas de compétences numériques.

Critères à étudier :

Le nombres de personnes utilisant internet sellons leurres Age.

Résultats et observations :

Ainsi, une personne de plus de 75 ans X sur deux n'a pas d'accès à Internet depuis son domicile (53%), alors que seuls 2% des 15-29 ans ne sont pas équipés.

C'est également le cas de 34% des personnes peu ou pas diplômés (contre 3% des diplômés du supérieur), et de 16% des ménages les plus modestes (contre 4% des ménages les plus aisés).

Les personnes vivant seules, en couple sans enfant, ou encore résidant dans les départements d'outre-mer sont également touchées par ce défaut d'équipement.

Mais utiliser Internet ne suffit pas pour maîtriser les fondamentaux du numérique. Ainsi, **38% des usagers apparaissent manquer d'une compétence numérique** dans au moins un de ces domaines : la recherche d'information, la communication, l'utilisation de logiciels et la résolution de problèmes. 2% ne savent pas utiliser un ordinateur, même s'ils ont l'équipement nécessaire.

Recommandations et décisions :

La réduction de cette fracture numérique. Cela passera par de la formation, distribués par Pôle emploi, la Caisse d'allocations familiales, l'Assurance maladie et les collectivités.