

TP 2 : AFC

1) 1^{ère} étape : Mise en place d'un programme pour une AFC

On commence donc par implémenter quelques fonctions qui vont nous servir, à commencer par une fonction transformant les données fournies en matrice de correspondance (fonction *tabl_corres*). On transforme ensuite cette matrice en matrice de fréquence, matrice sur laquelle on va pouvoir travailler matrice F .

Pour réaliser l'AFC, on s'occupe d'abord du profil ligne. On va diagonaliser la matrice A , définie comme suit :

$$A = D_p^{-1/2} F' D_n^{-1} F D_p^{-1/2}$$

(Les termes D_p et D_n sont définis comme dans le cours)

Après diagonalisation en base orthonormée, on enlève les valeurs et vecteurs propres associées à la valeur propre la plus haute, pour s'aligner avec les résultats obtenus avec le package *FactoMineR*. Cet axe principal pourrai correspondre à un axe trivial, ne permettant pas de tirer assez d'informations sur les données traitées.

On obtient les nouvelles coordonnées de nos données à l'aide de la formule :

$$\psi_\alpha = D_n^{-1} F' D_p^{-1} u_\alpha D_p^{1/2}$$

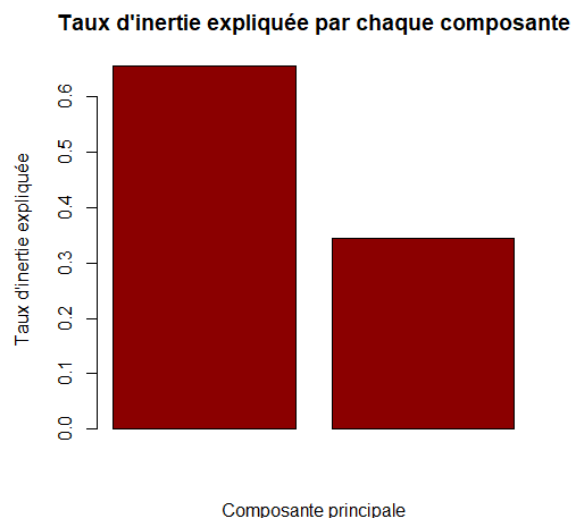
Il faut maintenant effectuer le même travail pour le profile ligne, pour cela, il suffit d'entrer non pas F mais F' en argument de notre fonction AFC.

Ainsi, en choisissant le nombre de composantes principales, on peut tracer nos profiles lignes et colonnes.

On implémente ensuite une fonction *Qualité* nous permettant de calculer la qualité de notre projection, comme en ACP.

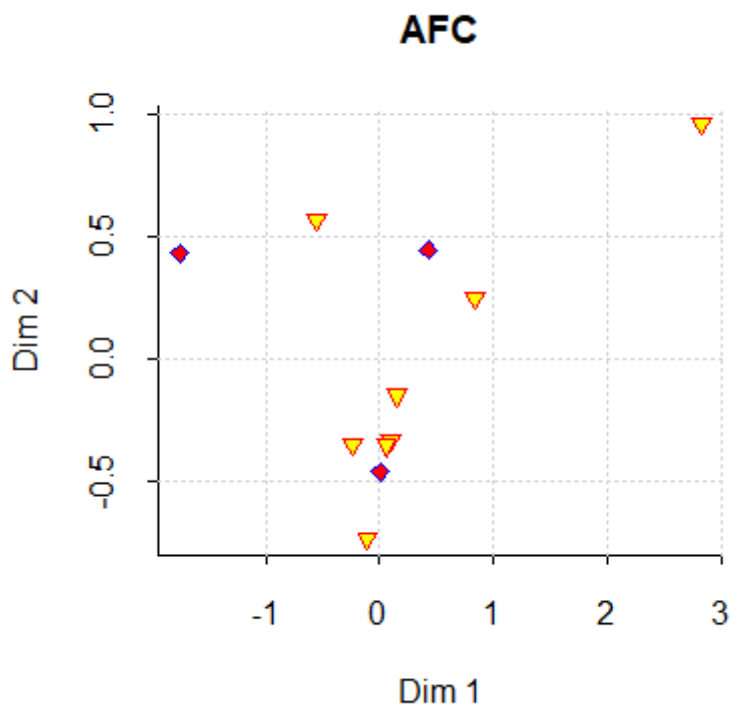
On va maintenant appliquer nos fonctions sur le fichier *TP_AFC_data.csv*, et comparer nos résultats avec ceux de la fonction CA.

On commence donc par étudier les valeurs propres de la matrice A , correspondant à l'inertie :

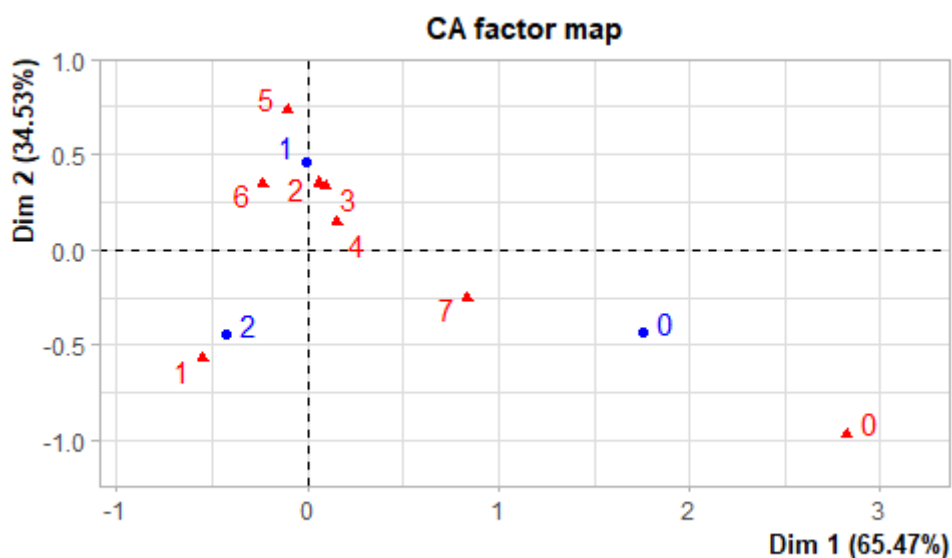


Ainsi, on voit qu'il n'y a que deux valeurs propres non nulles (le choix est donc rapide, même si toutefois dans le cas où l'une d'elle serait très grande par rapport à l'autre, on pourrait faire le choix de représenter nos données en une dimension).

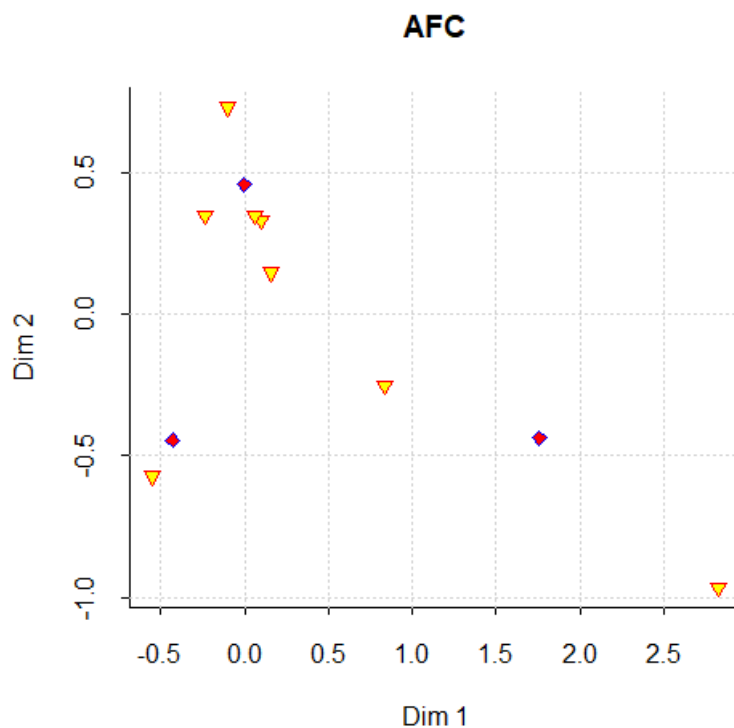
On réalise donc la projection, ce qui nous donne le résultat suivant :



Puisqu'il n'existe que deux composantes principales, il est inutile de s'intéresser à la qualité de la projection ; elle sera de 100%. On va maintenant comparer ces résultats avec de la fonction CA du package *FactoMineR*. Voici le résultat obtenu :



On peut penser que nos résultats sont différents de ceux de la fonction CA, mais après avoir regardé les coordonnées précises de la projection, on conclut qu'il ne s'agit en fait que des signes des vecteurs propres qui diffèrent (l'opposé d'un vecteur propre orthonormé est aussi un vecteur propre orthonormé). Pour s'en convaincre, on arrange les signes de nos coordonnées projetées pour obtenir la figure suivante :



On constate que nos résultats coïncident avec ceux de la fonction CA.