```
clear all; close all;
i=1;

%Données

data=  [68 76 127 60 72 68 84 68 82;
        2 4 3 3 2 3 2 2 4;
        81 74 41 81 101 102 111 68 70;
        36 34 31 37 35 40 42 44 24;
        64 115 172 82 60 76 83 57 237;
        89 53 69 84 64 33 30 38 57;
        25 93 87 13 66 6 3 14 45;
        6 12 20 4 4 8 8 7 22;
        7 3 1 12 10 6 4 12 1;
        24 33 25 21 28 23 23 13 20;
        58 38 24 41 41 26 32 51 18;
        84 90 80 136 87 14 14 9 13;
        85 91 84 135 89 134 187 159 64;
        6 7 2 3 8 5 11 6 1;
        12 18 13 13 10 6 3 10 8;
        17 15 11 11 14 14 14 14 12];

tab_1={ 'RFA' 3;
        'France' 6;
        'Italie' 6;
        'Pays Bas' 8;
        'Belg. Lux.' 10;
        'Russie' 6;
        'Irlande' 7;
        'Danemark' 8;
        'Grece' 5};

tab_2={ 'Cereales' 8;
        'Riz' 3;
        'Pommes de terre' 14;
        'Sucre blanc' 11;
        'Legumes' 7;
        'Fruits' 6;
        'Vin' 3;
        'Huiles vegetales' 16;
        'Margarine' 9;
        'Viande bov.' 11;
        'Viande porc.' 12;
        'Volailles' 9;
        'Lait et deriv.' 14;
        'Beurre' 6;
        'Fromages' 8;
        'Oeufs' 5 };

%-----------------------------------
traitement-----------------------------------%
```

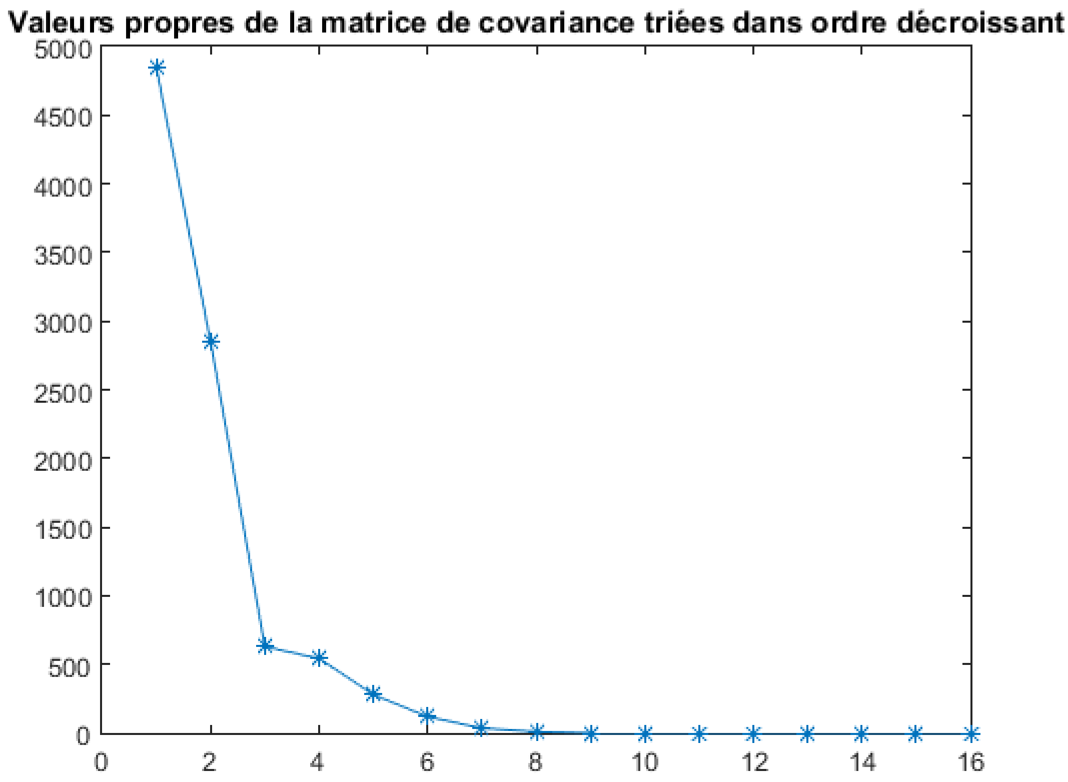
```
Y=[data(:,1),data(:,2),data(:,3),data(:,4),data(:,5),data(:,6),data(:,7),data(:,8)
Y=Y';                                   %On transpose les données
[n,m]=size(Y);                          %Récupération des dimensions de la
 matrice Y
Moyx=[];                                %Initialisation de la matrice des
 Moyennes
X=[];                                   %Initialisation de la matrice X


%Remplissage de la matrice des moyennes et de la matrice X
Moyx=mean(Y);                           %Moyenne de la matrice de données
X=[X, Y-ones(n,1)*Moyx];                %Matrice des données centrées
M=(1/n)*X'*X;                           %Matrice de covariance

[p,lambda]=(eig(M));                    % Vecteurs et valeurs propres de
 la matrice M
ValeursPropres=diag(lambda);            % On prend juste les valeurs
 propres qui sont situées sur la diagonale
Lambda=flipud(abs(ValeursPropres'));% On les classe par ordre
 décroissant et on prend la valeur absolue

figure(1);
plot(Lambda,'*-');                      %Affichage de l'évolution des
 valeurs propres
title('Valeurs propres de la matrice de covariance triées dans ordre
 décroissant');


%Axes factoriels constituant le repère de l'espace plus petit que le
 précédent

e=[];                                   %Initialisation de la matrice des
 axes factoriels

for i=1:n
    e(:,i)=p(:,i);                      %Les axes factoriels sont les
 vecteurs propres
end

%Taux d'inertie pour chaque axe factoriel

Taux=Lambda/sum(Lambda);
figure(2);
plot(Taux,'*-');
title('Taux inertie sur chaque axe factoriel');

%Taux d'inertie pour les 2 premiers axes factoriels
Taux1=Taux(1);
Taux2=Taux(2);

Xstar=X*p;                              %Matrice du nouveau nuage de
 points dans le nouvel espace défini par les axes factoriels
figure(3);
```

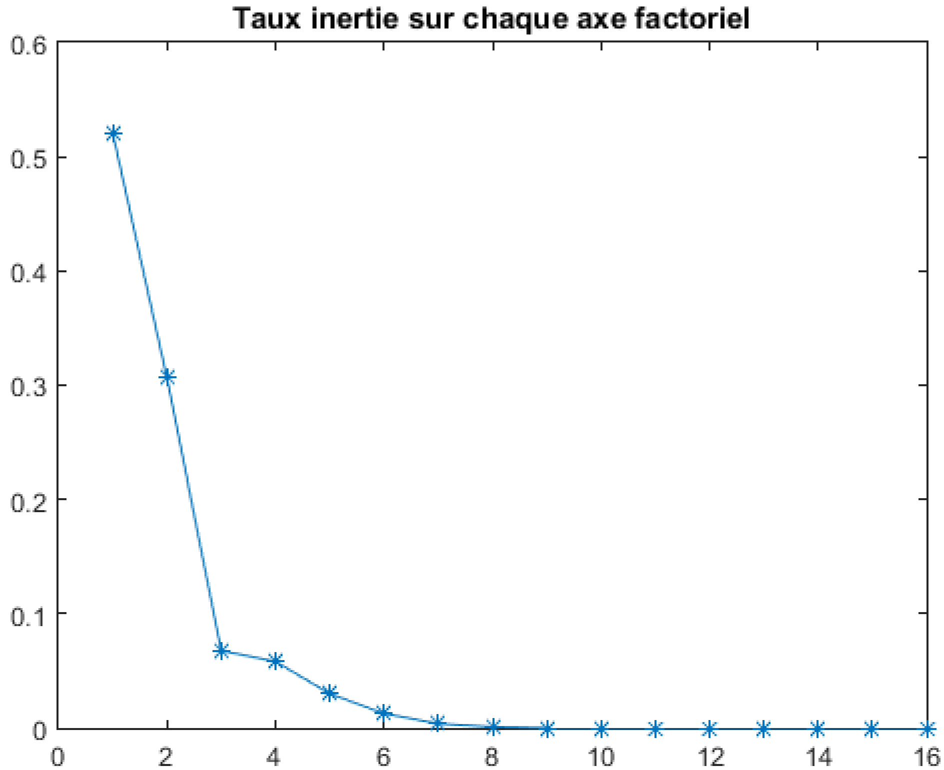
```
plot(Xstar(:,1),-Xstar(:,2),'*');   %Affichage du nouveau nuage de
 point sur les 2 premiers axes factoriels
title('Taux inertie sur les axes factoriels 1 et 2');

%A=cross(e(1:3,1),e(1:3,2));        %Produit vectoriel afin d'avoir
 les coefficients d'une équation du plan
                                    %On prend les 3 premières
 coordonnées des 2 axes factoriels principaux

%Affichage du nouveau nuage de points sur des axes factoriels
 différents des 2 premiers
figure(4);
plot(Xstar(:,4),-Xstar(:,5),'*');   %Axes factoriels 4 et 5
title('Taux inertie sur les axes factoriels 4 et 5');
figure(5);
plot(Xstar(:,7),-Xstar(:,8),'*');   %Axes factoriels 7 et 8
title('Taux inertie sur les axes factoriels 7 et 8');

%Cercle des corrélations

Ecarttype=[];                       %Initialisation de la matrice des
 écarts-types
figure(6);
hold on;

%--------------------Action sur les
 colonnes---------------------------%

e1=Xstar(:,1);                      %On définit l'axe factoriel 1
e2=Xstar(:,2);                      %On définit l'axe factoriel 2
Ecarttype=ones(n,1)*std(Y);         %On construit la matrice des
 ecarts-types
Z=X./Ecarttype;                     %On définit les données initiales
 réduites dans la matrice Z

rho1=( (Z')*(e1) )./( n*sqrt( Lambda(1) ) );
rho2=( (Z')*(e2) )./( n*sqrt( Lambda(2) ) );

plot(rho1,-rho2,'o');
text(rho1,-rho2,tab_2(:,1));

%Tracé du cercle centré en 0 et de rayon 1
t = 0:0.05:2*pi;
x = cos(t);
y = sin(t);
plot(x,y,'k-');                     %Affichage du cercle

plot([0,0],[-1,1],'-');
plot([-1,1],[0,0],'-');
axis('equal');
xlabel(['e1(',num2str(Taux1*100),'%)']);
ylabel(['e2(',num2str(Taux2*100),'%)']);
title('Cercle des corrélations');
```

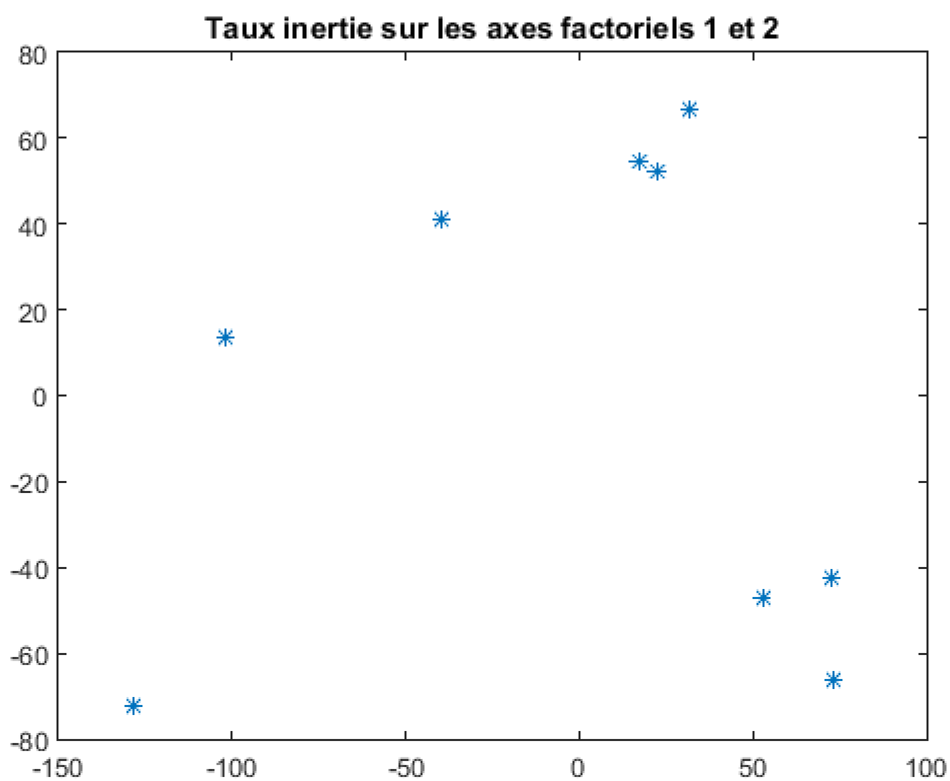
%-----------Synthèse------------%

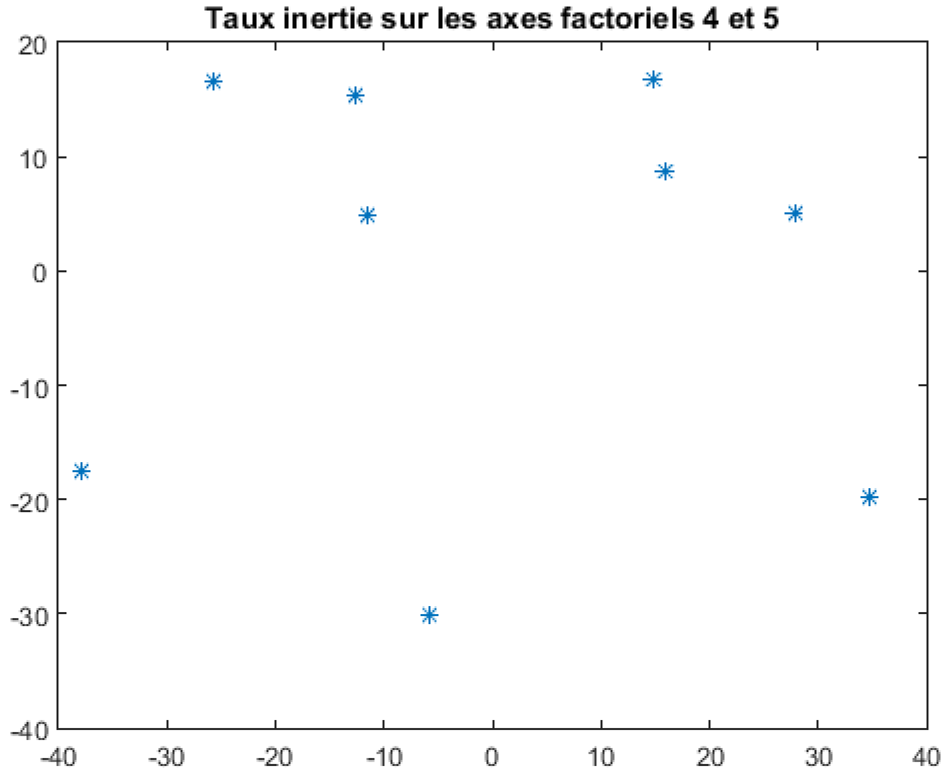
% L'objectif de cet exercice est d'utiliser la méthode ACP sur des données mettant en œuvre différentes habitudes alimentaires de plusieurs pays afin de savoir s'il existe un moyen de distinguer les pays selon leurs habitudes alimentaires.
% La méthode ACP consiste à transformer des variables liées entre elles en nouvelles variables décorrélées les unes des autres. Cela nous permet de réduire le nombre de variables ainsi que la dimension de l'espace représentée par les données de base. Les nouvelles variables sont appelées « composantes principales » et sont projetées sur des axes dits factoriels afin de représenter le nouvel espace.
% Dans notre cas, les données sont fournies par une matrice de 16 lignes et 9 colonnes d'entiers. Nous transposons cette matrice afin d'obtenir une ACP sur les produits et non sur les pays. Nous travaillons ensuite sur la matrice des données centrées en soustrayant à chaque valeur du tableau (ou de la matrice) la moyenne de ce dernier (de cette dernière).
% Nous déterminons la matrice de covariance de la matrice des données centrées et en récupérons les vecteurs propres et les valeurs propres. Cela nous permet de faire une première analyse sur les valeurs propres en constatant que ces dernières décroissent rapidement (lorsqu'on les classes de manière décroissante), mais surtout qu'à partir de la 8ème valeur propre, elles sont toutes proches de 0. Nous développerons cette analyse lorsque nous étudierons les taux d'inertie par axe factoriel. En récupérant les vecteurs propres, nous avons les axes factoriels de notre nouvel espace. En effet, les axes factoriels sont les vecteurs propres de la matrice de covariance, le premier correspondant au vecteur propre associé à la valeur propre la plus grande, le deuxième associé à la deuxième valeur propre la plus grande etc. Lorsque nous avons établis ces axes, nous travaillons sur la matrice du nouveau nuage de points dans le nouvel espace défini par ces axes. Nous pouvons maintenant étudier le taux d'inertie, suivant exactement la même allure de courbe que les valeurs propres, ainsi que le taux d'inertie entre 2 axes factoriels. C'est ce que nous avons fait entre les axes factoriels 1 et 2, 4 et 5 puis 7 et 8. Nous constatons alors que plus nous prenons un taux d'inertie entre deux axes factoriels grands, plus les valeurs du nuage de point associé et affiché se rapprochent de 0. Cela correspond tout à faire à l'analyse faite sur les valeurs propres qui se rapprochent de plus en plus de 0 (à partir de la 8ème valeur). Si nous avions pris entre les axes factoriels 11 et 12, le nuage de point aurait été extrêmement proche de 0.
% Ensuite, nous allons mettre en place la matrice des données initiales réduites. Pour cela, nous calculons l'écart-type des données initiales.
% Nous affichons enfin les variables réduites de la méthode ACP ainsi qu'un cercle de rayon 1 centré en 0 pour pouvoir visualiser les résultats. Pour cela, nous nous plaçons sur les deux premiers axes factoriels afin d'avoir des résultats plus visibles et exploitables. Ce cercle des corrélations montre alors la répartition des habitudes alimentaires non plus par pays mais de manière générale, avec comme effet une réduction de dimensionnalité. Nous passons alors d'un individu composé de variables à une simple variable Par ailleurs,

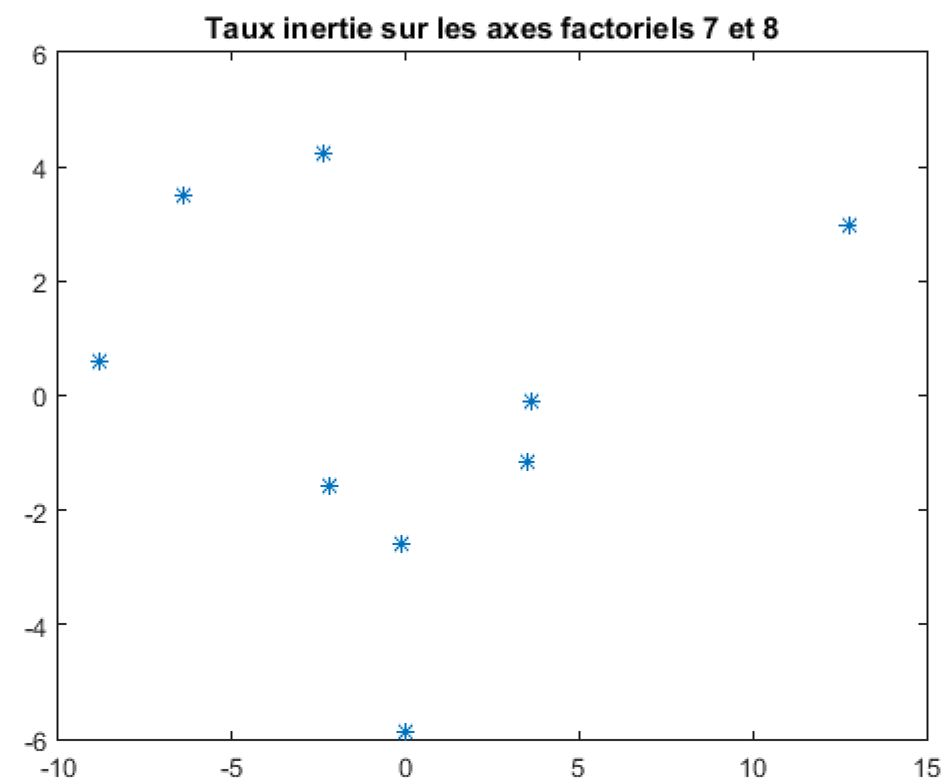
étant donné que nous sommes dans le cas d'une projection sur un plan factoriel, l'ensemble des points du nuage seront à l'intérieur du cercle unité. Nous pouvons alors, grâce au cercle dessiné, déterminer la corrélation de la variable. En effet, plus un point est proche de la circonférence du cercle, plus elle est corrélée avec l'une des deux composantes principales associées au plan factoriel.
En reprenant notre exemple, nous avons alors la variable réduite associée au beurre qui est très corrélé avec l'axe factoriel e1, alors que la variable réduite associée aux Volailles est très corrélée avec l'axe factoriel e2.
% Nous avons alors appris grâce à cet exercice à réduire la dimension d'un problème et exprimé un individu en une simple variable visualisable sur un graphique qui est exploitable grâce aux positions géographiques des points. Par ailleurs, nous pouvons changer d'axes factoriels afin d'obtenir différentes approches du problème, plus ou moins précises selon les vecteurs propres de la matrice de covariance. Le taux d'inertie et le rendu visuel final nous donnent également des informations sur le problème. Cette méthode est utilisable quelle que soit la dimension de la matrice de données de base, ce qui implique que cette dernière n'a pas forcément de limites pour l'appliquer.

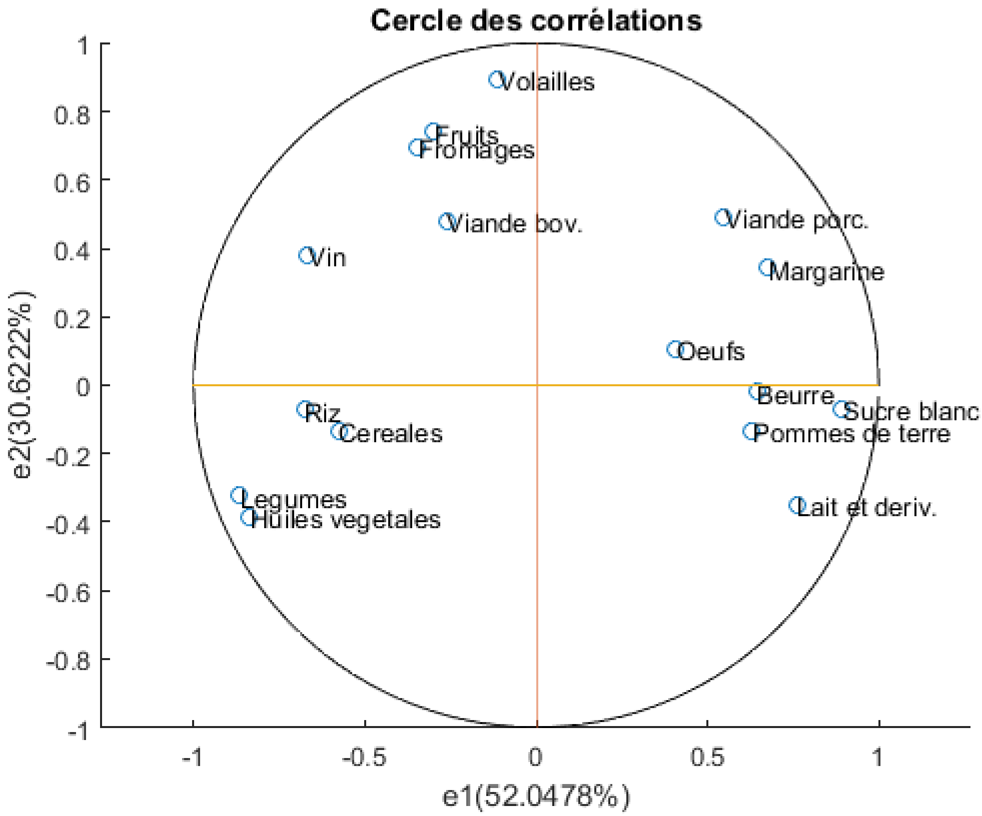



Taux inertie sur chaque axe factoriel
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10
12
14
16


Taux inertie sur les axes factoriels 1 et 2
80
60
40
20
0
-20
-40
-60
-80
-150
-100
-50
0
50
100


Taux inertie sur les axes factoriels 4 et 5
20
10
0
-10
-20
-30
-40
-40
-30
-20
-10
0
10
20
30
40


Taux inertie sur les axes factoriels 7 et 8
6
4
2
0
-2
-4
-6
-10
-5
0
5
10
15



Published with MATLAB® R2015a