

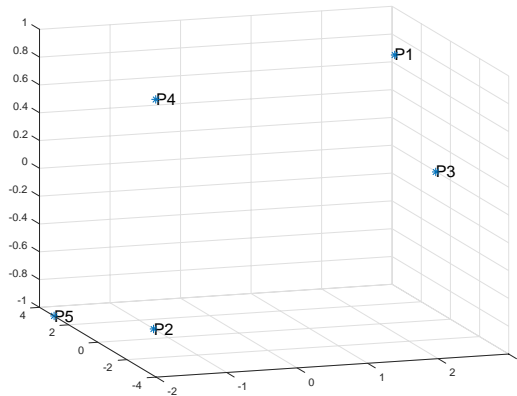
Analyse en Composantes Principales (ACP) Travaux Pratiques Matlab

Exercice 1 : un nuage de points presque coplanaires

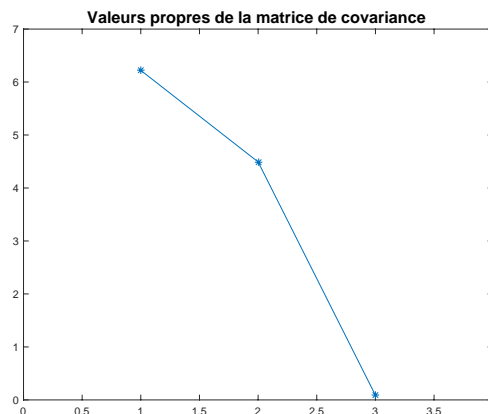
On considère un nuage de 5 points de \mathbb{R}^3 dont les coordonnées sont les suivantes :

$$P_1(2, -1, 1), \quad P_2(-1, 1, -1), \quad P_3(3, 1, 0), \quad P_4(-2, -4, 1), \quad P_5(-2, 3, -1),$$

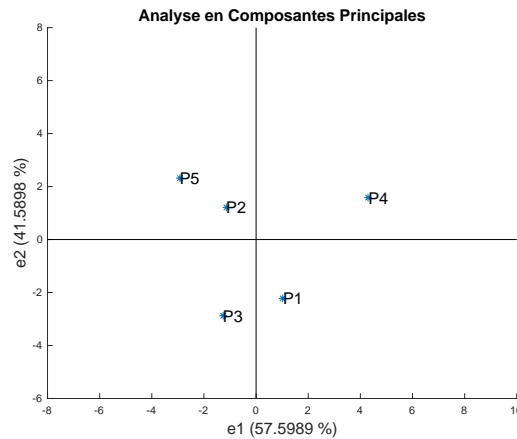
- 1) Afficher le nuage de points dans un repère à trois dimensions (commande `plot3`).
On constate, en faisant tourner le repère, que les points sont presque coplanaires.



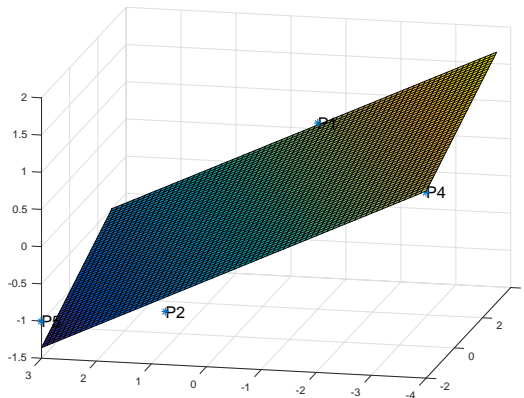
- 2) On souhaite alors déterminer l'équation du plan passant au plus proche des 5 points du nuage. Pour cela on propose d'abord le problème à l'aide de l'ACP (en reprenant les notations du cours) :
- Identifier les variables et les individus puis en déduire les valeurs de n et m
 - Construire la matrice des données (on la notera Y) et affecter les valeurs à n et m (commande `size`)
 - Construire la matrice des données centrée (on la notera X , utiliser la commande `mean`, pas de boucle `for`).
Que constatez-vous ?
 - Construire la matrice de covariance (on la notera M)
 - Diagonaliser la matrice de covariance.
 - Placer les valeurs propres dans un tableau, que l'on notera `lambda`, dans l'ordre croissant puis placer dans une matrice, que l'on notera P , les vecteurs propres en respectant l'ordre des valeurs propres (commandes `flipud` et `fliplr`).
 - Afficher la courbe des valeurs propres (comme ci-dessous)



- Calculer le taux d'inertie expliquée pour chaque axe factoriel.
- Afficher la matrice P . Que représentent les vecteurs colonnes de cette matrice ?
- Calculer et afficher la matrice des composantes principales (que l'on notera X_{star}).
- Afficher le nuage de points projeté sur les deux premier axes factoriels (comme ci-dessous)



- L'équation du plan défini par les deux premiers axes factoriels e_1 et e_2 a pour équation $ax + by + cz + d = 0$ où a, b et c sont les composantes d'un vecteur normal au plan factoriel (e_1, e_2) .
 - o Quelle est la valeur de d ?
 - o Calculer à l'aide de Matlab les valeurs de a, b et c
 - Indication : on peut fixer $a = 1$ et montrer que le problème revient à résoudre un système de la forme $A.U = V$ où $U = (b, c)^T$, A est une matrice et V un vecteur que l'on précisera
 - o Afficher, dans le repère 3D de la question 1, le plan d'équation $ax + by + cz + d = 0$ (commandes `quiver` et `surf`)



- 3) Exécuter à nouveau votre programme en remplaçant le nuage de points précédents par les points suivants et vérifier si tout fonctionne correctement :

$$P_1(2,1,3), \quad P_2(-2,5,-3), \quad P_3(3,-1,0), \quad P_4(5,-1,3), \quad P_5(-2,3,-1), \quad P_6(1,2,-4), \quad P_7(-4,-2,1)$$

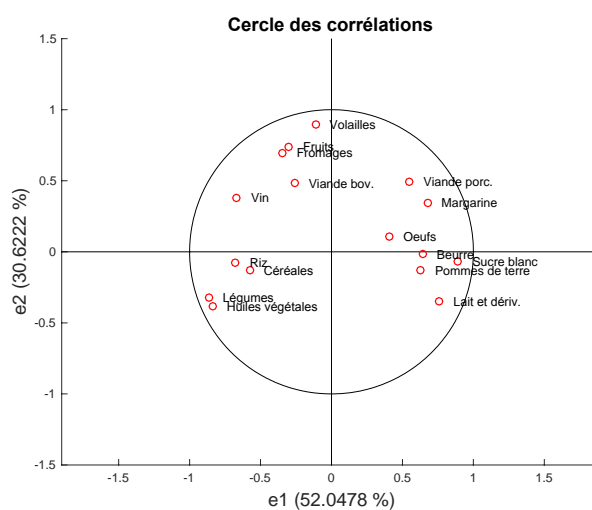
Exercice 2 : habitudes alimentaires

Utiliser l'analyse en composante principale des données présenter dans le tableau ci-dessous afin de voir s'il existe un moyen de distinguer les pays selon leurs habitudes alimentaires.

	RFA	France	Italie	Pays Bas	Belg. Lux.	Russie	Irlande	Danemark	Grèce
Céréales	68	76	127	60	72	68	84	68	82
Riz	2	4	3	3	2	3	2	2	4
Pommes de terre	81	74	41	81	101	102	111	68	70
Sucre blanc	36	34	31	37	35	40	42	44	24
Légumes	64	115	172	82	60	76	83	57	237
Fruits	89	53	69	84	64	33	30	38	57
Vin	25	93	87	13	66	6	3	14	45
Huiles végétales	6	12	20	4	4	8	8	7	22
Margarine	7	3	1	12	10	6	4	12	1
Viande bov.	24	33	25	21	28	23	23	13	20
Viande porc.	58	38	24	41	41	26	32	51	18
Volailles	84	90	80	136	87	14	14	9	13
Lait et dériv.	85	91	84	135	89	134	187	159	64
Beurre	6	7	2	3	8	5	11	6	1
Fromages	12	18	13	13	10	6	3	10	8
Œufs	17	15	11	11	14	14	14	14	12

Quelques indications :

- Récupérer le fichier Ex2_data.m sur CPe-campus
- S'inspirer fortement du travail effectué pour l'exercice 1, notamment pour les notations (qui sont celles du cours), les affichages, etc ...
- Mettre en évidence le taux d'inertie expliquée pour chaque axe factoriel.
- On pourra éventuellement ajouter des graphiques montrant la projection du nuage sur des plans autres que le plan factoriel principal (e_1, e_2) et interpréter ces résultats.
- Afficher le cercle des corrélations (cf. cours) sur le plan factoriel principal et interpréter les résultats (comme ci-dessous).

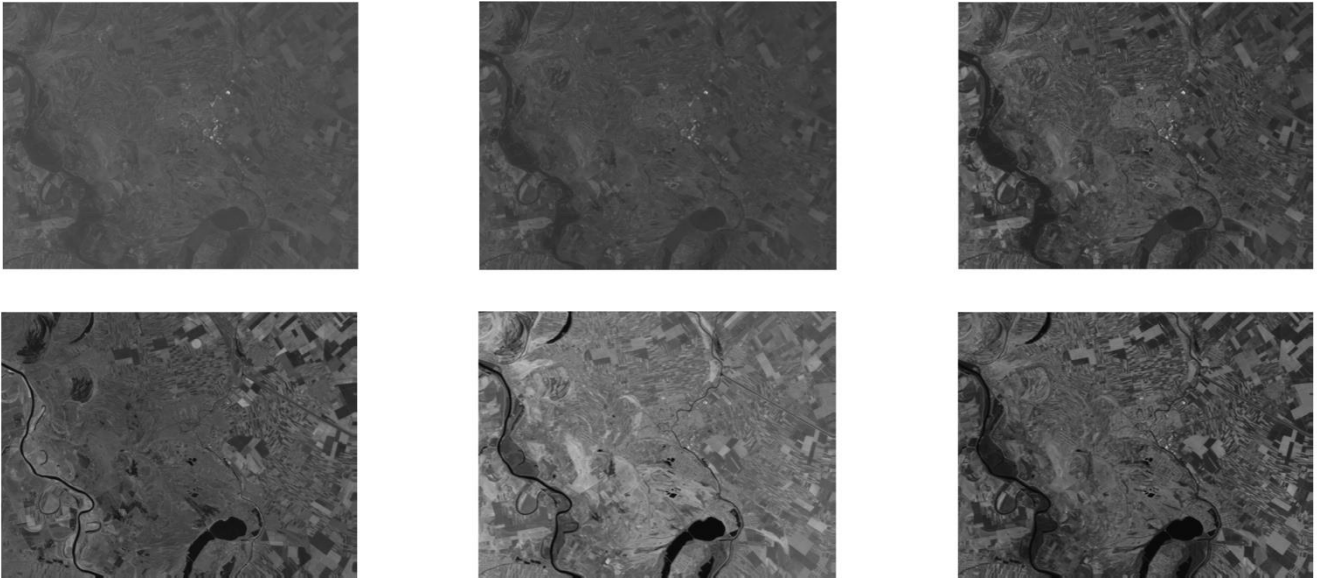


Une fois votre travail finalisé, ajoutez à la fin de votre code un paragraphe dans lequel vous faites une synthèse de votre travail (quel est le contexte, quels sont les objectifs, que fallait-il faire, pourquoi et comment on l'a fait, quel sont les résultats obtenus, qu'a-t-on appris d'important, ...), puis utilisez l'outil « Publish » de Matlab pour générer un document au **format pdf** que vous enverrez à serge.mazauric@cpe.fr avant le 05/03/2018 minuit. Le nom du fichier doit avoir la structure suivante :

Exercice 3 : Transformée de Karhunen-Loève (TKL) – d'après Stéphane Grousson, 2003

La transformée de Karhunen-Loève est le nom que l'on donne à l'analyse en composantes principales lorsque cette dernière est appliquée au traitement des images.

A partir de 6 images satellitaires d'une même vue mais acquises à travers des canaux fréquentiels différents, nous proposons ici de construire une seule image en conservant le maximum d'informations.



Quelques indications :

- Récupérer sur CPe-campus les fichiers des 6 images `i1.jpg`, ..., `i6.jpg` ainsi que le fichier Matlab `Ex3_tmp.m`
- Exécuter le programme `Ex3_tmp.m`
- Construire la matrice des données (que l'on notera Y) comportant 6 colonnes (1 par image) et n lignes où n est le nombre de pixels d'une image ($n = H \times W$)
- Faire les calculs de l'ACP (matrice des données centrée X , matrice de covariance M , calcul et affichage des valeurs propres dans l'ordre décroissant, matrice de projection P , matrice des composantes principales X_{star})
- Calculer l'inertie expliquée par le premier axe factoriel
- Construire la matrice (que l'on notera I_{rec}) de dimension $H \times W$ à partir de la première composante principale (commande `reshape`)
- Afficher l'image correspondant à la matrice I_{rec} et commenter.

Une fois votre travail finalisé, ajoutez à la fin de votre code un paragraphe dans lequel vous faites une synthèse de votre travail (quel est le contexte, quels sont les objectifs, que fallait-il faire, pourquoi et comment on l'a fait, quel sont les résultats obtenus, qu'a-t-on appris d'important, ...), puis utilisez l'outil « Publish » de Matlab pour générer un document au **format pdf** que vous enverrez à serge.mazauric@cpe.fr avant le 05/03/2018 minuit. Le nom du fichier doit avoir la structure suivante :

NOM1_NOM2_TP_ACP_EX3.pdf

(majuscules obligatoires)