

Projet : Matrices et ACP

Consignes :

- **Résoudre** les différents exercices à l'aide du **logiciel R sous R-studio**
- Écrire les **scripts** pour chaque traitement
- Présenter les résultats dans un **pdf** avec tous les tracés et tableaux de calcul
- **Commenter** et **interpréter** les résultats lors de la soutenance

Exercice 1

Voici la méthode de la puissance pour évaluer une valeur propre strictement dominante :

1. Choisir un vecteur initial x_0 dont la plus grande composante est 1
2. Pour $k = 0, 1, \dots$
 - a. Calculer Ax_k
 - b. Appeler μ_k une composante de Ax_k dont la valeur absolue est la plus grande possible
 - c. Calculer $x_{k+1} = \frac{1}{\mu_k} Ax_k$
3. La suite (μ_k) s'approche de la valeur propre dominante et la suite (x_k) s'approche d'un vecteur propre associé, à peu près quel que soit le choix de x_0

Calculer μ_k pour $k = 1, \dots, 4$ dans les cas suivants

1. $A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
2. $A = \begin{pmatrix} -3 & 2 \\ 2 & 2 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
3. Faire le test pour une matrice symétrique d'ordre 8 et x_0 à composantes positives dont la plus grande est 1.

Exercice 2

On considère le modèle d'emploi suivant. Chaque personne dans la population est soit employé soit sans emploi. Les deux états de ce modèle sont "être employé" et "être sans emploi". Soit $x_k = (a_k, b_k)$, avec a_k la fraction de population qui est employée à la fin de la période n et b_k la fraction sans-emploi. Supposons qu'une personne employée a une probabilité de 90% de rester employée à la période suivante et qu'une personne sans-emploi a une probabilité de 60% de rester sans-emploi à la période suivante.

1. Donner le taux de chômage à long terme.
2. Pour $x_0 = (20, 16)$, tracer x_k , pour les 15 premières années.

Exercice 3

Soit le système

$$(S) : \begin{cases} x' = 2y \\ y' = -2x - 4x^3 - \varepsilon y \end{cases}$$

1. Cas $\varepsilon = 4$

On suppose qu'autour du point d'équilibre $(0, 0)$ le système est approché par

$$(S_1) : \begin{cases} x' = 2y \\ y' = -2x - 4y \end{cases}$$

- Écrire le système en utilisant une matrice A et montrer qu'elle est diagonalisable
- Résoudre ce système localement à $(0, 0)$ avec comme condition initiale $x(0) = 1, y(0) = 1$. Tracer les solutions de (S_1)
- A l'aide des fonctionnalités de R, résoudre directement (S) , tracer les solutions et comparer avec la question 1.b

2. Cas $\varepsilon = 8$

On suppose qu'autour du point d'équilibre $(0, 0)$ le système est approché par

$$(S_2) : \begin{cases} x' = 2y \\ y' = -2x - 8y \end{cases}$$

- Écrire le système en utilisant une matrice A et montrer qu'elle est diagonalisable
- Résoudre ce système localement à $(0, 0)$ avec comme condition initiale $x(0) = 1, y(0) = 1$. Tracer les solutions de (S_2)
- A l'aide des fonctionnalités de R, résoudre directement (S) , tracer les solutions et comparer avec la question 2.b

Exercice 3 : véhicules vendus en France

On dispose de données techniques sur 52 véhicules vendus en France (année modèle 1994). Les variables sont : **Puiss** (puissance, en chevaux fiscaux), **Cyl** (cylindrée, en cm³), **Long** (longueur), **Larg** (largeur), **Surf** (surface) **Poids** (poids total en kg), **Vites** (vitesse maximum en km/h), **DepArr** (temps, en secondes, pour parcourir 1000m, départ arrêté), **Conso** (consommation moyenne aux 100km, en litres d'essence ou gazole). Les marques des véhicules sont abrégées **A1** (Alfa-Romeo), **BM** (BMW), **Ci** (Citroën), **Fi** (Fiat), **Pe** (Peugeot) et **Re** (Renault). Les données sont reproduites ci-dessous

	Puiss	Cyl	Long	Larg	Surf	Poids	Vites	DepArr	Conso		Puiss	Cyl	Long	Larg	Surf	Poids	Vites	DepArr	Conso
A13315ie	7	1490	4.07	1.61	6.58	970	180	31.4	7.7	FiTEMPRATD	6	1929	4.35	1.70	7.38	1190	178	34.0	6.0
A115520	10	1995	4.44	1.70	7.55	1215	205	31.3	8.0	FiCROMA20	10	1995	4.52	1.76	7.95	1250	190	32.1	8.9
A11557D	6	1929	4.44	1.70	7.55	1250	180	35.0	6.4	FiCROMA25TD	8	2500	4.52	1.76	7.95	1370	192	33.0	7.2
A116425T	7	2500	4.55	1.76	8.02	1490	202	32.2	6.8	Pe106XN	4	954	3.56	1.58	5.62	760	145	39.0	5.9
BM316i	7	1596	4.43	1.70	7.53	1190	191	34.1	7.9	Pe106XTD	4	1360	3.56	1.58	5.62	840	145	40.1	4.8
BM325TD	7	2498	4.43	1.70	7.53	1335	198	33.3	6.9	Pe306XN	6	1360	3.99	1.69	6.75	1020	165	36.2	6.7
BM518i	10	1796	4.72	1.75	8.26	1360	194	33.6	8.6	Pe306XTDT	6	1905	3.99	1.69	6.75	1120	180	34.2	6.0
BM730i	16	2986	4.91	1.85	9.06	1600	222	30.1	11.1	Pe405GL	6	1360	4.41	1.69	7.46	1030	169	36.3	7.1
CiAX10	12	954	3.53	1.55	5.48	706	149	38.3	5.8	Pe405SR	7	1761	4.41	1.69	7.46	1100	185	33.5	7.5
CiAX14i	7	1360	3.53	1.55	5.48	772	176	33.6	6.7	Pe405STD	6	1905	4.41	1.69	7.46	1150	181	34.4	7.2
CiAX14TD	4	1360	3.53	1.55	5.48	720	152	38.3	4.2	Pe605SLi	9	1998	4.72	1.80	8.50	1295	199	33.2	9.5
CiZX11	5	1124	4.07	1.70	6.93	935	161	37.6	6.6	Pe605Sv24	16	2975	4.72	1.80	8.50	1460	235	28.6	9.5
CiZX16	7	1580	4.07	1.70	6.93	995	177	34.9	7.5	Pe605SLD	6	2138	4.72	1.80	8.50	1360	176	36.5	6.6
CiZX19TD	5	1905	4.07	1.70	6.93	1050	185	34.0	5.9	ReTwingo	5	1239	3.43	1.63	5.59	790	150	35.9	6.5
CiXANTIA16I	7	1580	4.44	1.75	7.80	1124	175	36.4	7.9	ReClio12RN	5	1171	3.71	1.62	6.02	825	155	38.8	6.2
CiXANTIA18I	9	1761	4.44	1.75	7.80	1176	187	34.3	7.9	ReClio14RN	6	1390	3.71	1.62	6.02	850	175	33.0	6.4
CiXANTIA20i	11	1998	4.44	1.75	7.80	1238	198	33.3	8.3	ReClio19D	6	1870	3.71	1.62	6.02	905	161	36.0	5.5
CiXANTIA1TD	6	1905	4.44	1.75	7.80	1252	180	35.5	6.5	Re1914RN	7	1390	4.16	1.70	7.08	965	173	34.0	6.6
CiXM20i	11	1998	4.71	1.79	8.45	1310	201	33.3	9.5	Re1918RN	7	1794	4.16	1.70	7.08	1045	176	33.6	7.7
CiXNV6	16	2963	4.71	1.79	8.45	1420	222	33.3	9.5	Re1919dT	5	1870	4.16	1.70	7.08	1080	183	32.9	5.9
CiXMTD12	6	2088	4.71	1.79	8.45	1396	192	34.6	6.4	Re21PrimaTS	7	1721	4.53	1.73	7.82	1010	185	32.1	7.5
FiUNO14ie	6	1372	3.69	1.56	5.75	860	165	34.0	7.4	Re21PrimaTD	6	2068	4.53	1.73	7.82	1130	177	33.5	6.1
FiUNOTD	5	1367	3.69	1.56	5.75	910	165	34.0	5.3	ReSafrane20i	9	1995	4.73	1.82	8.61	1370	189	33.8	8.8
FiTIP018ie	9	1756	3.96	1.70	6.73	1180	183	33.0	8.8	ReSafrane25d	6	2499	4.73	1.82	8.61	1565	195	34.3	7.6
FiTIP0TD	6	1929	3.96	1.70	6.73	1150	175	34.0	5.9	ReEspace22i	11	2165	4.43	2.06	9.13	1320	175	33.9	9.7
FiTEMPRA18ie	9	1756	4.35	1.70	7.38	1190	188	32.7	8.5	ReEspace21dT	6	2068	4.43	2.06	9.13	1350	162	36.5	7.5

A. Analyse rapide des variables

On donne ci-dessous la matrice des corrélations des variables

	Puiss	Cyl	Long	Larg	Surf	Poids	Vites	DepArr	Conso
Puiss	1.0000	0.5911	0.5073	0.4098	0.5093	0.5236	0.7239	-0.5722	0.8005
Cyl	0.5911	1.0000	0.7566	0.6417	0.7611	0.8853	0.8385	-0.6379	0.5412
Long	0.5073	0.7566	1.0000	0.7436	0.9574	0.9078	0.7853	-0.5388	0.6646
Larg	0.4098	0.6417	0.7436	1.0000	0.9039	0.7799	0.4700	-0.3064	0.6091
Surf	0.5093	0.7611	0.9574	0.9039	1.0000	0.9154	0.7048	-0.4713	0.6897
Poids	0.5236	0.8853	0.9078	0.7799	0.9154	1.0000	0.8031	-0.5558	0.6366
Vites	0.7239	0.8385	0.7853	0.4700	0.7048	0.8031	1.0000	-0.8306	0.7136
DepArr	-0.5722	-0.6379	-0.5388	-0.3064	-0.4713	-0.5558	-0.8306	1.0000	-0.6171
Conso	0.8005	0.5412	0.6646	0.6091	0.6897	0.6366	0.7136	-0.6171	1.0000

1. Quelles sont les couples de variables particulièrement peu corrélées entre elles ?
2. Pour chacune des 4 premières composantes principales, donner la liste des individus qui contribuent à l'axe de manière significative.

B. Analyse des composantes principales

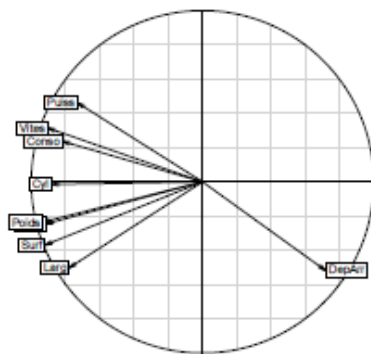
On procède à une analyse en composantes principales sur variables centrées-réduites de ces données. On donne ci-dessous les écarts types des 5 premières composantes principales.

2.5404	1.0666	0.8105	0.5762	0.4867
--------	--------	--------	--------	--------

3. Expliquez leur lien avec les valeurs propres des axes principaux et déduisez-en le nombre d'axes principaux que l'on doit garder pour l'analyse. Quelle part d'inertie totale sera alors représentée ?

On donne ci-dessous, pour les deux premiers axes seulement, les corrélations des variables (ainsi que le cercle des corrélations), les coordonnées des individus, leurs contributions aux axes (en 10000e) et leur qualités de représentation par les axes (en 10000e encore). Les tableaux concernant les individus sont coupés en deux pour des raisons de hauteur.

	Comp1	Comp2
Puiss	-0.7290	0.4609
Cyl	-0.8804	-0.0143
Long	-0.9128	-0.2450
Larg	-0.7777	-0.5015
Surf	-0.9185	-0.3689
Poids	-0.9317	-0.2348
Vites	-0.9027	0.3142
DepArr	0.7176	-0.5175
Conso	-0.8168	0.2361



	Axis1	Axis2
Al3315ie	0.8182	1.6240
Al15520	-1.7575	1.3451
Al155TD	-0.0320	-0.7715
Al16425T	-2.2647	-0.3321
BM316i	-0.4492	0.0770
BM325TD	-1.3575	0.0357
BM518i	-2.1277	-0.0048
BM730i	-6.3269	1.4291
CiAX10	4.1120	1.2849
CiAX14i	2.8777	2.0598
CiAX14TD	4.9161	-0.1716
CiZX11	2.3793	-1.0183
CiZX16	0.8765	0.2116
CiZX19TD	0.8538	-0.0884
CiXANTIA16I	0.0208	-0.9455
CiXANTIA18I	-0.9011	-0.0410
CiXANTIA20i	-1.8132	0.6566
CiXANTIA2D	-0.2124	-1.1799
CiXM20i	-2.8558	0.2723
CiXNV6	-4.6574	1.2313
CiXMTD12	-1.4923	-1.5075
FiUNO14ie	2.6568	1.4433
FiUNOTD	3.1709	0.9036
FiTIP018ie	-0.2553	1.1851
FiTIP0TD	0.9294	-0.0634
FiTEMPRA18ie	-0.9106	0.8332

	Axis1	Axis2
FiTEMPRATD	0.2083	-0.4785
FiCROMA20	-2.0257	0.5872
FiCROMA25TD	-1.9179	-0.2922
Pe106XN	4.8125	-0.3927
Pe106XTD	4.7802	-0.9042
Pe306XN	1.8475	-0.4146
Pe306XTDT	0.8995	0.0395
Pe405GL	1.0516	-0.7955
Pe405SR	-0.2215	0.2153
Pe405STD	-0.0461	-0.3018
Pe605SLi	-2.6595	-0.0957
Pe605Sv24	-5.6430	2.3718
Pe605SLD	-1.0229	-2.1840
ReTwingo	3.7924	0.4409
ReClcio12RN	3.7648	-0.4618
ReClcio14RN	2.2830	1.2953
ReClcio19D	2.7024	0.1956
Re1914RN	1.1040	0.1368
Re1918RN	0.3078	0.3598
Re1919dT	0.5938	-0.0058
Re21PrimaTS	-0.5773	0.2439
Re21PrimaTD	-0.2688	-0.6991
ReSafrane20i	-2.4553	-0.7071
ReSafrane25d	-2.6130	-1.5805
ReEspace22i	-3.2612	-1.5574
ReEspace21dT	-1.6338	-3.4837

	Axis1	Axis2
Al3315ie	20	446
Al15520	92	306
Al155TD	0	101
Al16425T	153	19
BM316i	6	1
BM325TD	55	0
BM518i	135	0
BM730i	1193	345
CiAX10	504	279
CiAX14i	247	717
CiAX14TD	720	5
CiZX11	169	175
CiZX16	23	8
CiZX19TD	22	1
CiXANTIA16I	0	151
CiXANTIA18I	24	0
CiXANTIA20i	98	73
CiXANTIA2D	1	235
CiXM20i	243	13
CiXNV6	646	256
CiXMTD12	66	384
FiUNO14ie	210	352
FiUNOTD	300	138
FiTIP018ie	2	237
FiTIP0TD	26	1
FiTEMPRA18ie	25	117

	Axis1	Axis2
FiTEMPRATD	1	39
FiCROMA20	122	58
FiCROMA25TD	110	14
Pe106XN	690	26
Pe106XTD	681	138
Pe306XN	102	29
Pe306XTDT	24	0
Pe405GL	33	107
Pe405SR	1	8
Pe405STD	0	15
Pe605SLi	211	2
Pe605Sv24	949	951
Pe605SLD	31	806
ReTwingo	429	33
ReClcio12RN	422	36
ReClcio14RN	155	284
ReClcio19D	218	6
Re1914RN	36	3
Re1918RN	3	22
Re1919dT	11	0
Re21PrimaTS	10	10
Re21PrimaTD	2	83
ReSafrane20i	180	85
ReSafrane25d	203	422
ReEspace22i	317	410
ReEspace21dT	80	2051

	Axis1	Axis2
Al3315ie	1422	5601
Al15520	5867	3437
Al155TD	7	4266
Al16425T	6215	134
BM316i	1742	51
BM325TD	4253	3
BM518i	7879	0
BM730i	9247	472
CiAX10	6572	642
CiAX14i	6478	3319
CiAX14TD	9538	12
CiZX11	7319	1341
CiZX16	6564	383
CiZX19TD	3122	34
CiXANTIA16I	2	4247
CiXANTIA18I	6500	13
CiXANTIA20i	7992	1048
CiXANTIA2D	256	7907
CiXM20i	8629	78
CiXNV6	7914	553
CiXMTD12	3818	3896
FiUNO14ie	7145	2108
FiUNOTD	8447	686
FiTIP018ie	231	4967
FiTIP0TD	3841	18
FiTEMPRA18ie	4157	3480

	Axis1	Axis2
FiTEMPRATD	318	1678
FiCROMA20	8278	696
FiCROMA25TD	7339	170
Pe106XN	9558	64
Pe106XTD	9033	323
Pe306XN	8806	444
Pe306XTDT	4275	8
Pe405GL	3740	2140
Pe405SR	880	831
Pe405STD	37	1584
Pe605SLi	8567	11
Pe605Sv24	8255	1458
Pe605SLD	1625	7406
ReTwingo	9131	123
ReClcio12RN	9175	138
ReClcio14RN	6723	2164
ReClcio19D	8297	43
Re1914RN	6107	94
Re1918RN	1600	2187
Re1919dT	1484	0
Re21PrimaTS	1452	259
Re21PrimaTD	326	2206
ReSafrane20i	8439	700
ReSafrane25d	6313	2310
ReEspace22i	5145	1173
ReEspace21dT	1471	6689

4. Quelles variables déterminent les deux premières composantes principales (précisez les critères utilisés) ?
5. Que peut-on dire d'un éventuel effet de taille ? Expliquez ce qui se passe (et comment arranger les choses).
6. Comment peut-on interpréter la première composante principale ? Et la seconde ?
7. Quels sont les individus qui déterminent les deux premiers axes principaux ? (précisez les critères utilisés).
8. Quels sont les 5 individus les moins bien représentés par le sous espace qu'on a déterminé en question 3 ?

C. Influence du type de moteur

On s'intéresse aux véhicules à moteur diesel (les autres ont un moteur à essence). On en détermine la liste en regardant les modèles de voitures dont le nom contient la lettre "D" ou "d". Les coordonnées du barycentre des véhicules diesel sur les deux premiers axes principaux sont respectivement 0.4451 et -0.6599 .

9. Combien de véhicules correspondent à ce critère ? Expliquez comment ce critère définit une variable supplémentaire, et comment on l'utilise.
10. Calculer les valeurs test de la modalité "diesel" sur les deux premiers axes. Peut-on en déduire une nouvelle interprétation des axes principaux ? On supposera le nombre d'individus suffisant.