

Analyse Factorielle

Pr Roch Giorgi

 roch.giorgi@univ-amu.fr

Objectif

- Étudier simultanément un nombre important de variables quantitatives
 - ✓ 2 variables quantitatives \Rightarrow nuage de points (espace de dimension 2)
 - ✓ n variables \Rightarrow espace de dimension n !
- Méthode pour obtenir un résumé « pertinent » des données initiales
 - ✓ Travail sur les intercorrélations entre les variables pour en extraire des dimensions plus globales
 - ✓ Revenir à un espace de dimension réduite

Différences avec l'ACP

- Porte sur la proportion de variance que l'on cherche à expliquer
 - ✓ ACP
 - Variance totale présente dans la matrice de corrélation
 - ✓ AF
 - Variance commune partagée par certaines variables
- L'objectif principal d'une AF n'est pas toujours une réduction des données
 - ✓ Mise en évidence de la structure latente des données

Principe (1)

- Intercorrélations dans les données dues à l'effet de facteurs communs

$$V_i = \hat{a}_i^1 F_1 + \hat{a}_i^2 F_2 + \dots + U_i$$

✓ $\hat{a}_i^1, \hat{a}_i^2, \dots$

- Déterminés par l'analyse
- Différents d'une variable à l'autre

✓ F_1, F_2, \dots

- Facteurs communs aux différentes variables (V_i)
- Non directement observables
- Définis par différents regroupement de variables

✓ U_i : variance unique de la $i^{\text{ème}}$ variable

- Variance spécifique (non soumise aux facteurs)
- Variance d'erreur

Principe (2)

- Utilisation de la variance commune
 - ✓ On ne travail plus sur la matrice de corrélations classique
 - ✓ Travail sur la matrice de corrélations réduite
 - La diagonale contient l'estimation de la variance commune
 - La variance totale de chacune des variables $\neq 1$

	AF	ACP	
Prix	0,738	1	Variance commune 73,8% Variance spécifique 26,2% (négligeable)
Quantité	0,912	1	
Alcool	0,866	1	Variance commune 49,9% Variance spécifique 51,1% (élevé)
Prestige	0,499	1	
Couleur	0,922	1	
Arôme	0,857	1	
Goût	0,881	1	
Total	5,675	7	

L'AF étudiera 81,1% de la variance commune

Principe (3)

- Extractions des facteurs
 - ✓ Processus itératif
 - ✓ Calcul de la variance commune de chaque variable à chaque étape
 - ✓ Ancienne variance commune remplacée par la nouvelle
 - ✓ Arrêt quand le changement de variance commune suffisamment faible

Matrice de Corrélation

- Ne doit pas être singulière
 - ✓ Déterminant $> 0,00001$
- Ne doit pas être la matrice d'identité
 - ✓ Déterminant $\neq 1$
 - ✓ Test de sphéricité de Barlett
- Intercorrélation pour toutes les variables
 - ✓ Inspection visuelle de la matrice de corrélation
 - ✓ Mesure d'adéquation de l'échantillonnage de Kaiser-Meyer-Olkin

Extraction des Facteurs

- Critère de Kaiser
- Test d'accumulation de variance de Cattell (scree test)
- Analyse parallèle de Horn
- Facilité à interpréter les facteurs à extraire
- Confrontation des méthodes pour décider

Matrice des Facteurs

- Donne les poids factoriels de chacune des variables
- Extraction des variables sur chacun des facteurs
- Permet une représentation graphique

Rotation des Axes

- Techniques de rotations axiales comme pour ACP
 - ✓ Rotation orthogonale
 - ✓ Rotation oblique
- Modifie la matrice des poids factoriels

Remarques

- Comparabilité des solutions entre l'AF et l'ACP dans certaines situations
 - ✓ Nombre de variables > 30
 - ✓ Variance commune d'aucune des variables $< 0,4$
- Choix fonction de l'objectif
 - ✓ AF
 - Expliquer la variance commune entre certaines variables à l'aide de variables latentes
 - ✓ ACP
 - Réduire une masse de données en conservant le maximum de variance totale de ses variables

Exemple de Présentation

- Enquête sur les consommateurs de bière sur leurs motivations d'achat (Wuensch, 2005)
- n=220
- 7 variables (cotées sur une échelle de 1 à 100)
 - ✓ Coût faible d'un pack de 6 bières
 - ✓ Quantité de la bouteille
 - ✓ Haut degré d'alcool
 - ✓ Prestige de la marque
 - ✓ Couleur de la bière
 - ✓ Arôme
 - ✓ Goût

Matrice de Corrélation

Matrice de corrélation

	Prix	Quantité	Alcool	Prestige	Couleur	Arôme	Goût
Corrélation Prix	1,000	,832	,767	-,406	,018	-,046	-,064
Quantité	,832	1,000	,904	-,392	,179	,098	,026
Alcool	,767	,904	1,000	-,463	,072	,044	,012
Prestige	-,406	-,392	-,463	1,000	-,372	-,443	-,443
Couleur	,018	,179	,072	-,372	1,000	,909	,903
Arôme	-,046	,098	,044	-,443	,909	1,000	,870
Goût	-,064	,026	,012	-,443	,903	,870	1,000

Déterminant : 0,001

Test de sphéricité de Bartlett : $p < 0,001$

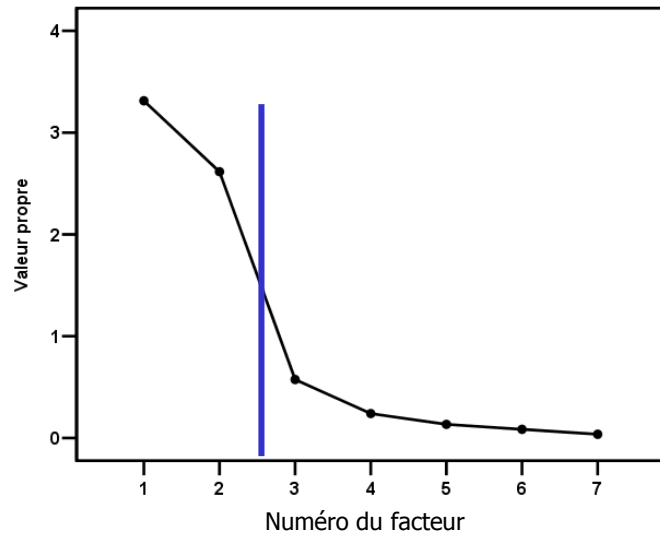
	Indice d'adéquation de KMO
Couleur	0,779
Arôme	0,550
Prestige	0,630
Goût	0,763
Prix	0,590
Alcool	0,801
Quantité	0,676
Global	0,665

Extraction des Facteurs (1)

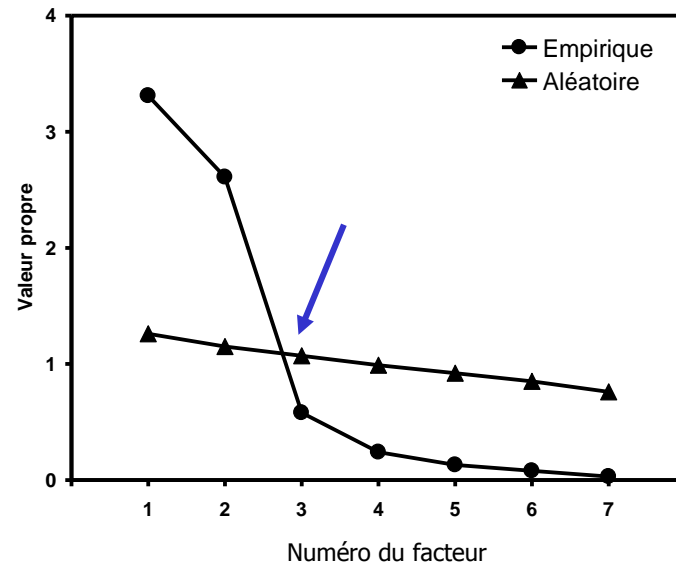
Facteurs	Valeurs propres initiales		
	Total	% de la variance	% cumulés
1	3,313	47,327	47,327
2	2,616	37,369	84,696
3	0,575	8,209	92,905
4	0,24	3,427	96,332
5	0,134	1,921	98,252
6	0,085	1,221	99,473
7	0,037	0,527	100
Total	7	100	



Extraction des Facteurs (2)



Test d'accumulation de variance de Cattell (scree test)

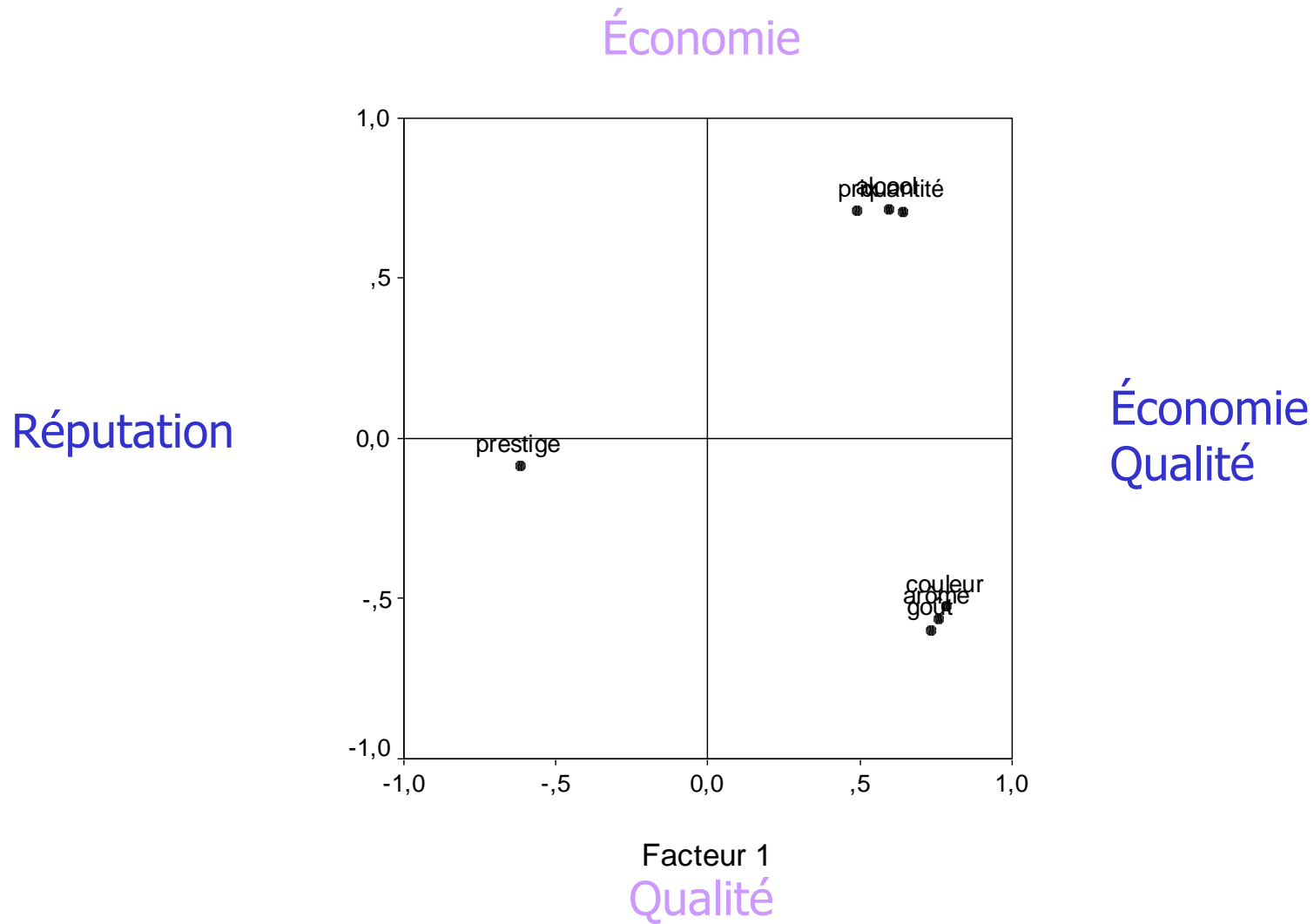


Analyse parallèle de Horn

Matrice des Facteurs

	Facteur	
	1	2
Couleur	0,785	-0,526
Arôme	0,759	-0,565
Goût	0,735	-0,601
Prestige	-0,614	-0,088
Alcool	0,595	0,715
Prix	0,494	0,708
Quantité	0,644	0,706

Graphe des Facteurs



Rotation des Axes : VARIMAX (1)

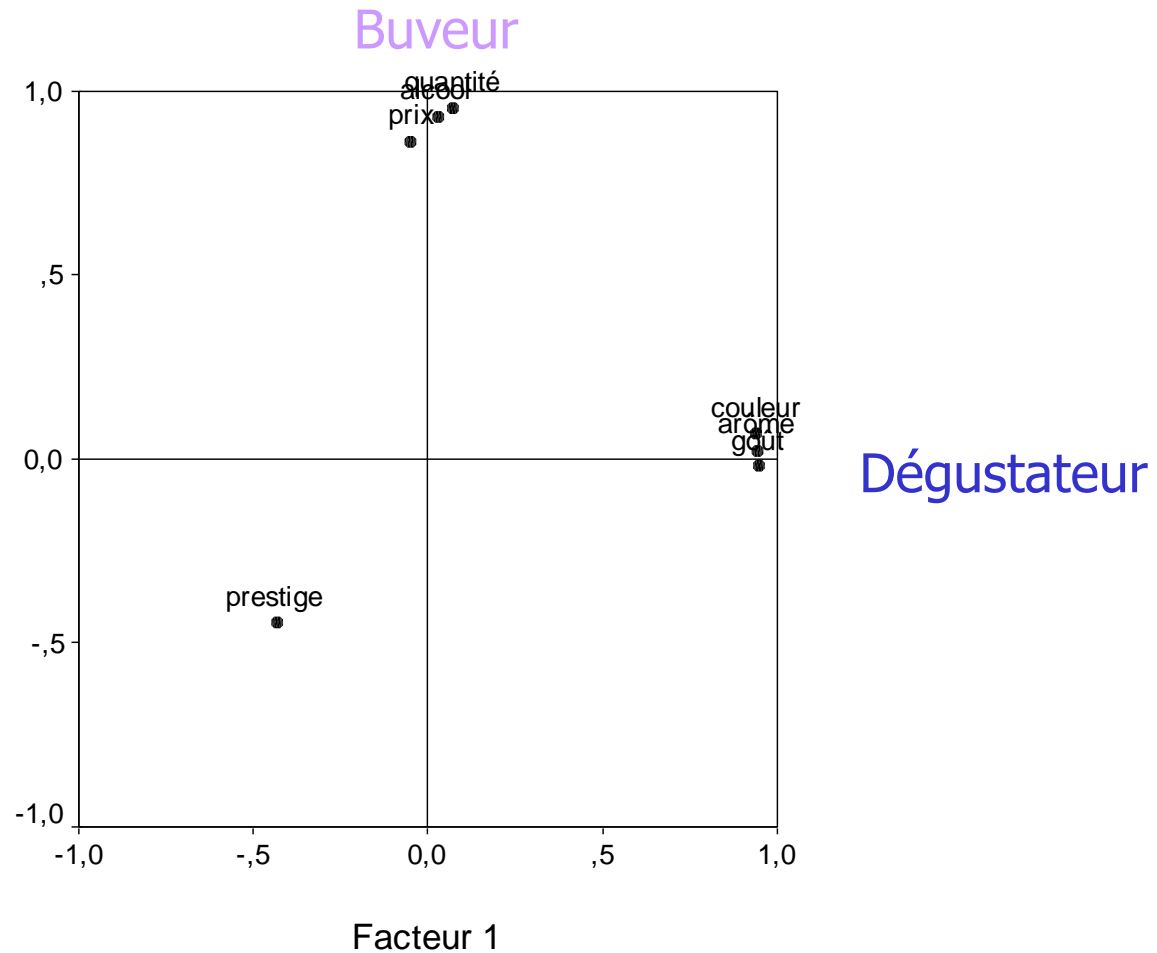
	Facteur	
	1	2
Couleur	0,785	-0,526
Arôme	0,759	-0,565
Goût	0,735	-0,601
Prestige	-0,614	-0,088
Alcool	0,595	0,715
Prix	0,494	0,708
Quantité	0,644	0,706

Sans rotation

	Facteur	
	1	2
Goût	0,950	-0,022
Arôme	0,946	0,021
Couleur	0,942	0,068
Quantité	0,073	0,953
Alcool	0,030	0,930
Prix	-0,046	0,862
Prestige	-0,431	-0,447

Rotation VARIMAX

Rotation des Axes : VARIMAX (2)



Source

- Baillargeon J. L'analyse factorielle exploratoire (2003)
<http://www.uqtr.ca/cours/srp-6020/afe/afe.pdf>