

Tema 4. APLICACIONES LINEALES

ANÁLISIS DE COMPONENTES PRINCIPALES (Principal Components Analysis, PCA)

Ejercicios: ENUNCIADOS

1. Análisis y valoración de grandes superficies (FUOC, PCA, 2.4 Procedimientos/Actividad 2.4, pág. 45-46 (pág. 12a13 del doc))

Hemos realizado un estudio acerca de diferentes aspectos relativos a diez grandes superficies de venta y hemos valorado de 0 a 10 las cuestiones siguientes:

- Facilidad de acceso al complejo comercial (A).
- Comodidad de aparcamiento (B).
- Bondad de los precios (C).
- Calidad de los productos (D).
- Servicio de guardería (E).
- Ofertas y promociones de artículos (F).
- Variedad de productos (G).
- Área de descanso y de ocio (H).

Hemos llegado a los siguientes resultados valorativos:

- Cuadro de puntuaciones (que nos dan como datos de partida):

Aspecto valorado / Gran Superf	A	B	C	D	E	F	G	H
#1	6	9	8	7	6	7	9	6
#2	7	9	5	3	7	5	4	7
#3	6	8	5	3	7	4	4	6
#4	5	5	7	2	4	4	3	4
#5	6	7	9	9	7	8	9	5
#6	8	9	9	7	7	7	8	8
#7	2	1	5	6	2	4	5	3
#8	7	8	3	2	6	2	3	6
#9	4	3	8	6	2	9	7	4
#10	5	6	7	7	4	8	8	6

Se desea representar e interpretar dichos resultados para saber lo que más valoran los clientes y así tomar decisiones sobre estrategias de mejora de los mismos.

A partir de la resolución matemática que se muestra más abajo, basada en el análisis de componentes principales, responder a las siguientes cuestiones:

- 1.1 Identificar variables de observación, elementos observados y tipos de datos (valores indicados) a tratar.
- 1.2 ¿Qué se consigue con el método de resolución propuesto?.
- 1.3 En este caso en concreto, qué buscamos y cómo nos va a ayudar a resolver el problema.
- 1.4 ¿Qué pasos sigue la resolución matemática propuesta?. ¿Es un resultado razonable / tiene sentido?
- 1.5 ¿Qué grandes superficies tiene más impacto sobre los productos y cuáles sobre los servicios.

Resolución matemática a interpretar:

Si calculamos la correlación entre todas estas características, comprobaremos que hay variables muy correlacionadas entre sí y que, por tanto, hay mucha información redundante:

- Correlaciones entre variables:

	A	B	C	D	E	F	G	H
A	1							
B	0,941	1						
C	0,077	0,040	1					
D	-0,136	-0,097	0,751	1				
E	0,882	0,926	0,022	-0,052	1			
F	-0,051	-0,053	0,846	0,805	-0,132	1		
G	0,053	0,115	0,812	0,949	0,067	0,868	1	
H	0,903	0,894	0,073	0,000	0,792	0,064	0,18	1

- Autovalores:

Valor propio	3,7031	3,5150	0,3285	0,2535	0,0958	0,0608	0,0422	0,0013
Proporción	0,463	0,439	0,041	0,032	0,012	0,008	0,005	0,000
Acumulativa	0,463	0,902	0,943	0,975	0,987	0,995	1,000	1,000

- Autovectores:

Variable	CP1	CP2
Acceso	-0,489	-0,132
Aparcamiento	-0,496	-0,124
Precios	-0,136	0,463
Calidad	-0,072	0,496
Guardería	-0,471	-0,134
Promoción	-0,088	0,494
Variedad	-0,169	0,490
Descanso	-0,483	-0,073

2. Análisis y valoración de ratios financieros (FUOC, PCA, 2.4 Procedimientos/Actividad 2.2, pág. 39-40 (pág. 6a7 del doc))

Suponemos tres ratios financieras calculadas sobre cinco cajas de ahorros:

Caja	Ratio 1 (X_1)	Ratio 2 (X_2)	Ratio 3 (X_3)
A	23	22	45
B	45	38	74
C	34	24	47
D	19	7	15
E	52	44	83

- 2.1 Calcular las correlaciones que se dan entre las tres variables y obtened los valores propios λ_1 , λ_2 y λ_3 .
- 2.2 Comprobar cómo el primer factor es capaz de absorber prácticamente el 98% de la información total, lo cual justifica que calculemos sólo un componente principal Z_1 .
- 2.3 Obtener el vector característico asociado a Z_1 .
- 2.4 Proyectar las cinco cajas de ahorros sobre el nuevo eje y discutir la ordenación conseguida (recordar que es necesario sustituir x_1 , x_2 y x_3 por los valores estandarizados).