

TAREA MA34B

Primavera 2007

Profesor: Alexis Peña

Auxiliar: Diego Díaz Espinoza

Condiciones:

- La tarea debe ser entregada el día jueves 22 de Noviembre del 2007 en secretaría de matemática (cuarto piso, saliendo del ascensor a la derecha y luego a la derecha pasando la mampara; tercera puerta). La secretaría está abierta desde las 9 de la mañana hasta las 17 horas.
- Se descontará 1 punto por cada día de atraso, esto implica:

Entrega	Máxima nota
Viernes	6.0
Sábado (en el examen)	5.0
Lunes	4.0
etc	etc

- Puede resolver la tarea en cualquier editor matemático – si así lo desea - (Maple, Matlab, otros) pero deberá entregar el desarrollo completo en hojas escritas a mano o en computador. Si sólo entrega resultados sin desarrollo se considerará con nota 1.0 la pregunta.
- No se aceptan copias. La tarea es individual. Cualquier intento de copia será sancionado con nota mínima y obligado a rendir el examen.
- Todas las preguntas valen los mismo.

1.- Sea X la siguiente matriz:

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

Consideremos los seis vectores $M = \{x_1, x_2, \dots, x_6\}$ dotado de la métrica euclidiana usual cuyas componentes están dadas por las de la matriz X.

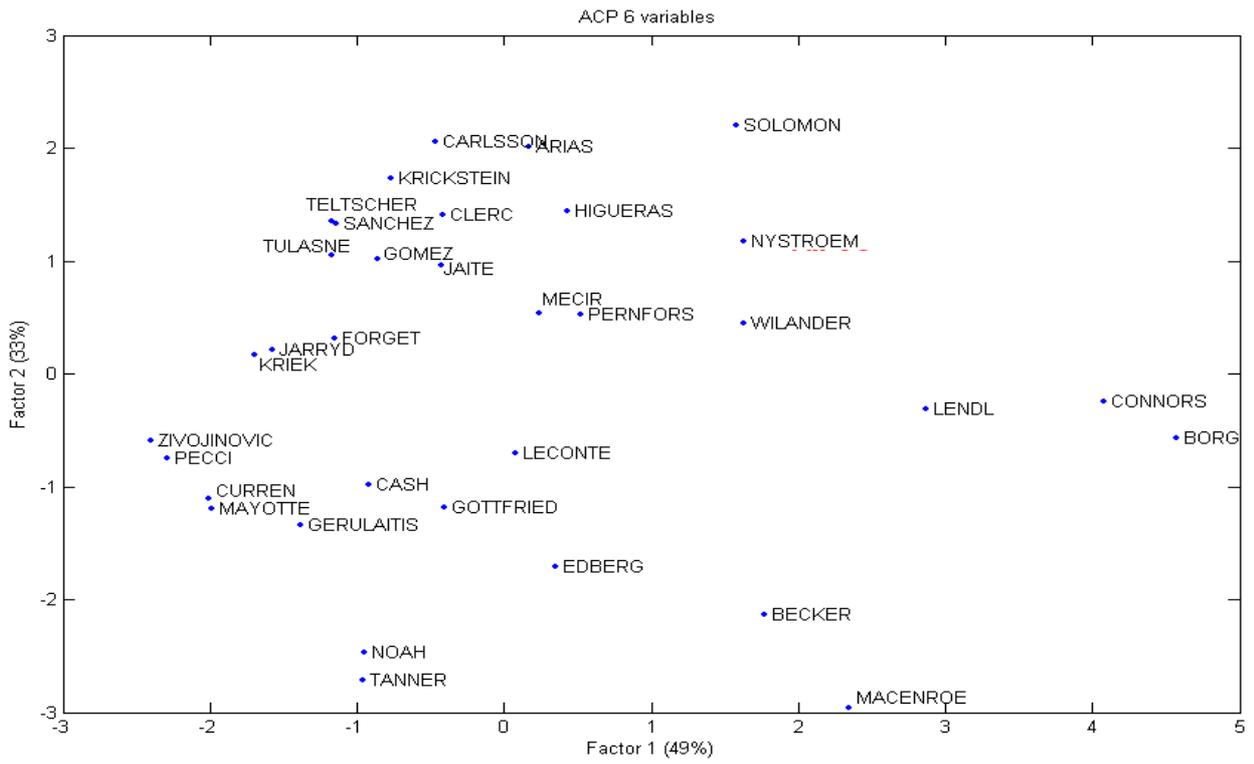
- Muestre que la nube de los puntos en R^3 está centrada en el origen. Calcule V y V^t .
- Determine los valores y vectores propios de V y V^t . Exponga la relación que existe entre los valores propios de ambas matrices (V y V^t).
- Calcule la traza de V y V^t .

2.- Se estudian 6 características de juego de 35 jugadores de tenis (derecho, revés, servicio, volea, retorno del servicio y el estado psíquico), que corresponden a notas entre 0 y 10. Los resultados del análisis en componentes principales sobre estos datos se encuentran en la tabla 1 y gráfico 1.

- a) Interprete los valores propios adjuntos (tabla 1). Dé las proporciones de la varianza reproducida por cada componente principal. Exprese la primera componente principal en función de las 6 variables y comente. ¿Pueden expresar las 6 variables a partir de las componentes principales? ¿Cómo?
- b) A partir de las correlaciones adjuntas (tabla 1), haga un gráfico de las variables sobre las 2 primeras componentes principales. Interprete el gráfico. Deduzca los coeficientes de correlación aproximados entre las 6 variables.
- c) Interprete el gráfico 1. En particular, en que difieren Connors, Pecci, Solomon y Mc Enroe.
- d) Un nuevo jugador VILAS tiene como valores (centradas y reducidas) para las 6 variables: 1.1418 0.8038 -0.6381 -0.9754 0.8507 0.9692. Cuales son sus coordenadas en el plano principal de los 2 primeros factores. Describe su juego y ¿a quien se parece su juego?
- e) Se quiere hacer la regresión lineal de una nueva variable el "Smash" sobre las dos primeras componentes principales sabiendo que sus correlaciones con las dos primeras C. P. son 0.1982 y -0.9022. Dé los coeficientes de la regresión sabiendo que la desviación estándar de la variable "Smash" es 1.9462.
- f) Dé el coeficiente de correlación múltiple. ¿Este último aproxima bien el coeficiente de correlación múltiple de "Smash" sobre las 6 otras variables?

Tabla 1: Correlaciones entre variables antiguas y componentes principales

	Componentes principales					
	1	2	3	4	5	6
Valor propio	2.935	1.9598	0.4557	0.3547	0.1657	0.1290
Derecho	0.7962	0.2659	-0.4014	0.3596	-0.0687	-0.0134
Revés	0.9162	-0.0525	-0.07	-0.3095	-0.0821	0.2246
Servicio	0.0104	-0.9448	-0.1438	0.1124	0.2533	0.0993
Volea	-0.1033	-0.9422	-0.1379	-0.0799	-0.2490	-0.1189
Retorno servicio	0.9348	-0.0042	-0.0095	-0.2175	0.1569	-0.2328
Estado psíquico	0.7597	-0.3254	0.5	0.2515	-0.0594	0.0120



3.- Se hizo un estudio sobre el nivel de desarrollo poblacional en el uso de las tecnologías de información, para lo cual se recolectó la siguiente información para 10 países:

Países	Area (Km2)	Población (millones)	Tasa Desempleo	Computadores por persona
E.E.U.U.	9629091	278,1	4,0	5,40
Namibia	825418	1,8	35,0	0,03
Francia	547030	59,6	9,7	0,89
Luxemburgo	2586	0,4	2,7	0,66
Finlandia	337030	5,2	9,8	1,20
Laos	236800	5,6	5,7	0,01
Chile	756950	15,3	8,0	0,56
Zimbawe	390580	11,4	50,0	0,03
Japan	377835	126,8	4,7	6,20
Kenia	582650	30,8	50,0	0,02

Para el estudio se utilizó un Análisis de Componentes Principales (ACP), del cual se presentan resultados en las tablas 5, 6 y 7:

Tabla 5

	Media	Desviación Estándar
Area	1368597,000	2763092,637
Población	53,504	83,497
Desempleo	17,960	18,252
Computadores	1,500	2,193

Tabla 6: Matriz de correlaciones

	Área	Población	Desempleo	Computadores
Área	1	0,895	-0,222	0,581
Población	0,895	1	-0,324	0,849
Desempleo	-0,222	-0,324	1	-0,471
Computadores	0,581	0,849	-0,471	1

Tabla 7: Valores y vectores propios normalizados

	λ1	λ2	λ3	λ4
Valor propio	2,753	0,874	0,353	0,020
	U1	U2	U3	U4
Área	0,517	0,393	-0,595	-0,473
Población	0,584	0,236	0,047	0,775
Desempleo	-0,318	0,881	0,346	-0,050
Computadores	0,539	-0,113	0,724	-0,415

- 3.1 Explique en qué consiste el análisis de componentes principales y cuáles son sus principales aplicaciones.
- 3.2 Interprete los resultados de las tablas 5 y 7.
- 3.3 Dibuje el círculo de correlaciones para los dos primeros ejes principales e interprete el resultado. Explique por qué es conveniente quedarse solo con estos.
- 3.4 Encuentre las componentes principales de cada país asociadas a los primeros ejes principales y grafique aproximadamente. Interprete los resultados.
- 3.5 Determine las contribuciones porcentuales de cada variable original en la construcción del eje.