Conteste todas las preguntas en el espacio asignado para las mismas. Si le falta espacio use la parte de atrás de la hoja.

Nombre completo: ____

1. Una serie de tiempo consiste en observaciones tomadas a lo largo del tiempo, usualmente a intervalos regulares. Estas observaciones suelen estar correlacionadas. Supongamos que tenemos las observaciones Z_1, Z_2, Z_3, Z_4 y que la correlación entre Z_t y Z_{t-k} es $corr(Z_t, Z_{t-k}) = \phi^k$ para todo t, donde $-1 < \phi < 1$ (Por ejemplo, $cor(Z_3, Z_1) = \phi^2$). Suponga que la media de las cuatro observaciones es 0 y la varianza de las cuatro es 1, y que el vector de las cuatro observaciones sigue una distribución normal multivariada.
   (a) (20 puntos) Determine la media condicional de Z_1 y Z_4 dados Z_2 y Z_3.

   (b) (20 puntos) la correlación parcial entre Z_t y Z_{t-k} se define como la correlación condicional entre Z_t y Z_{t-k} dadas las observaciones intermedias $Z_{t-1}, \ldots, Z_{t-k+1}$, es decir, $corr(Z_t, Z_{t-k}|Z_{t-1}, \ldots, Z_{t-k+1})$. Demuestre que la correlación parcial entre Z_4 y Z_1 es cero, es decir, $corr(Z_4, Z_1|Z_3, Z_2) = 0$.

2. Suponga que $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_p \end{bmatrix}^T$ es un vector aleatorio con vector de medias $\boldsymbol{\mu}$ y matriz de varianzas y covarianzas $\boldsymbol{\Sigma}$. Deseamos encontrar el primer componente principal, es decir, la combinación lineal $Y_1 = c_1 X_1 + \cdots + c_p X_p$ que tenga mayor varianza, tal que la norma del vector de coeficientes sea 1, es decir, $||\mathbf{c}|| = 1$, o en términos del cuadrado de la norma, $||\mathbf{c}||^2 = \boldsymbol{c}^T \boldsymbol{c} = c_1^2 + \cdots + c_p^2 = 1$.

   (a) (20 puntos) Demuestre que el vector de coeficientes $\boldsymbol{c}$ que maximiza la varianza de Y_1 sujeto a la restricción que $\boldsymbol{c}^T \boldsymbol{c} = 1$ es un vector propio de $\boldsymbol{\Sigma}$ (Sugerencia: utilice multiplicadores de Lagrange).

   (b) (20 puntos) Demuestre que la varianza del primer componente principal es el valor propio más grande de $\boldsymbol{\Sigma}$.

3. (20 puntos) Utilice el método k-means para separar las siguientes observaciones en dos conglomerados. Como conglomerados iniciales utilice la primera y segunda mitad de las observaciones.

X_1	X_2	X_3
1,89	1,83	1,89
1,21	1,17	1,21
3,08	2,89	3,08
2,66	2,93	2,66
2,64	1,81	2,64
1,43	1,55	1,43
4,53	1,34	4,53
-0,4	2,12	-0,4