Tema: Distancias, Repaso Algebra, Vectores y Matrices Aleatorios

1. Las siguientes son 5 medidas sobre las variables x_1, x_2, y x_3 :

x_1	9	2	6	5	8
x_2	12	8	6	4	10
x_3	3	4	0	2	1

   Encuentre las matrices $\overline{\mathbf{x}}, \mathbf{S}_n$, y $\mathbf{R}$.

2. Las 10 empresas más grandes a nivel mundial producen los siguientes datos:

Las 10 empresas más grandes a nivel mundial:

Company	$x_1 =$ sales (billions)	$x_2 =$ profits (billions)	$x_3 =$ assets (billions)
Citigroup	108.28	17.05	1,484.10
General Electric	152.36	16.59	750.33
American Intl Group	95.04	10.91	766.42
Bank of America	65.45	14.14	1,110.46
HSBC Group	62.97	9.52	1,031.29
ExxonMobil	263.99	25.33	195.26
Royal Dutch/Shell	265.19	18.54	193.83
BP	285.06	15.73	191.11
ING Group	92.01	8.10	1,175.16
Toyota Motor	165.68	11.13	211.15

   (a) Grafique el diagrama de dispersión para las variables x_1 y x_2. Comente la apariencia del diagrama.

   (b) Calcule $\bar{x}_1, \bar{x}_2, s_{11}, s_{22}, s_{12}$, y r_{12}. Interprete r_{12}.

   (c) Grafique los diagramas de dispersión para (x_2, x_3) y (x_1, x_3). Comente acerca de los patrones en ambas gráficas.

   (d) Calcule las matrices $\overline{\mathbf{x}}, \mathbf{S}_n$, y $\mathbf{R}$ para (x_1, x_2, x_3).

3. Dados los siguientes pares de medidas sobre dos variables x_1 y x_2:

x_1	-6	-3	-2	1	2	5	6	8
x_2	-2	-3	1	-1	2	1	5	3

   (a) Grafique los datos como un diagrama de dispersión y calcule s_{11}, s_{22} y s_{12}.

   (b) Usando $\tilde{x}_1 = x_1 \cos(\theta) + x_2 \sin(\theta)$ y $\tilde{x}_2 = -x_1 \sin(\theta) + x_2 \cos(\theta)$, calcule las medidas correspondientes sobre las variables $\widetilde{x}_1$ y $\widetilde{x}_2$, asumiendo que los ejes coordenados originales están rotados un ángulo de $\theta = 26$ grados.

   (c) Usando las medidas $\tilde{x}_1$ y $\tilde{x}_2$ de (b), calcule las varianzas de muestra $\widetilde{s}_{11}$ y $\widetilde{s}_{22}$

(d) Considere el nuevo par de medidas $(x_1, x_2) = (4, -2)$. Transforme estas medidas en $\tilde{x}_1$ y $\tilde{x}_2$ como en (b) y calcule la distancia $d(O, P)$ del nuevo punto $P = (\widetilde{x}_1, \tilde{x}_2)$ desde el origen $O = (0, 0)$, usando $d(O, P) = \sqrt{\frac{\widetilde{x}_1^2}{\widetilde{s}_{11}} + \frac{\widetilde{x}_2^2}{\widetilde{s}_{22}}}$. Nota: Necesitará $\widetilde{s}_{11}$ y $\widetilde{s}_{22}$ de (c).

(e) Calcule la distancia desde $P = (4, -2)$ hasta el origen $O = (0, 0)$ usando $d(O, P) = \sqrt{a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2}$ y las expresiones para a_{11}, a_{22}, y a_{12} de la siguiente diapositiva. Nota: necesitará s_{11}, s_{22}, y s_{12} de (a). Compare la distancia calculada aquí con la distancia calculada usando los valores $\widetilde{x}_1$ y $\widetilde{x}_2$ en (d). (Dentro del error de redondeo, los números deben ser los mismos).

Nota: para este ejercicio necesitara las expresiones de los coeficientes de la distancia estadística que incluye variabilidad y correlación:

La distancia desde $P = (\widetilde{x}_1, \widetilde{x}_2)$ hasta el origen $O = (0, 0)$ se puede escribir en términos de las coordenadas originales x_1 y x_2 de P cómo:

$$d(O, P) = \sqrt{a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2}$$

donde los coeficientes están dados por:

$$a_{11} = \frac{\cos^2(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} + \frac{\sin^2(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

$$a_{22} = \frac{\sin^2(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} + \frac{\cos^2(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

$$a_{12} = \frac{\cos(\theta)\sin(\theta)}{\cos^2(\theta)s_{11} + 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{22}} - \frac{\sin(\theta)\cos(\theta)}{\cos^2(\theta)s_{22} - 2\sin(\theta)\cos(\theta)s_{12} + \sin^2(\theta)s_{11}}$$

4. Sea $\mathbf{x}' = [5, \quad 1, \quad 3]$ y $\mathbf{y}' = \begin{bmatrix} -1, & 3, & 1 \end{bmatrix}$

   (a) Grafique los dos vectores.

   (b) Encuentre (i) la longitud de $\mathbf{x}$, (ii) el ángulo entre $\mathbf{x}$ y $\mathbf{y}$, y (iii) la proyección de $\mathbf{y}$ en $\mathbf{x}$.

   (c) Dado que $\bar{x} = 3$ y $\bar{y} = 1$, grafique $[5-3, 1-3, 3-3] = [2, -2, 0]$ y $[-1-1, 3-1, 1-1] = [-2, 2, 0]$

5. Cuando $\mathbf{A}^{-1}$ y $\mathbf{B}^{-1}$ existan, demuestre cada uno de los siguiente ítems:

   (a) $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$

   (b) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

   Sugerencia: la parte (a) se puede demostar si se observa que $\mathbf{AA}^{-1} = \mathbf{I}, \mathbf{I} = \mathbf{I}'$, y $(\mathbf{AA}^{-1})' = (\mathbf{A}^{-1})'\mathbf{A}'$. La parte (b) sigue de $(\mathbf{B}^{-1}\mathbf{A}^{-1})\mathbf{AB} = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$

6. Sea

$$\mathbf{A} = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$$

   (a) ¿Es $\mathbf{A}$ una matriz simétrica?
   (b) Muestre que $\mathbf{A}$ es definida positiva.
   (c) Determine los valores y vectores propios de $\mathbf{A}$.
   (d) Encuentre la descomposición espectral de $\mathbf{A}$.
   (e) Determine la inversa de $\mathbf{A}$.
   (f) Encuentre los valores y vectores propios de $\mathbf{A}^{-1}$.

7. Considere los conjuntos de puntos (x_1, x_2) cuyas "distancias" desde el origen están dadas por

$$c^2 = 4x_1^2 + 3x_2^2 - 2\sqrt{2}x_1x_2$$

   para $c^2 = 1$ y para $c^2 = 4$. Determine los ejes mayor y menor de las elipses de distancias constantes y sus longitudes asociadas. Dibuje las elipses de distancias constantes y comente sus posiciones. ¿Qué pasará cuando c^2 aumente?

8. Verifique las relaciones $\mathbf{V}^{1/2}\boldsymbol{\rho}\mathbf{V}^{1/2} = \mathbf{\Sigma}$ y $\boldsymbol{\rho} = \left(\mathbf{V}^{1/2}\right)^{-1}\mathbf{\Sigma}\left(\mathbf{V}^{1/2}\right)^{-1}$, donde $\mathbf{\Sigma}$ es el $p \times p$ matriz de covarianza poblacional, $\boldsymbol{\rho}$ es la matriz de correlación poblacional $p \times p$ y $\mathbf{V}^{1/2}$ es la matriz de desviación estándar de la población.

9. Derive las expresiones para la media y las varianzas de las siguientes combinaciones lineales en términos de las medias y covarianzas de las variables aleatorias $X_1, X_2,$ y X_3.

   (a) $X_1 - 2X_2$
   (b) $-X_1 + 3X_2$
   (c) $X_1 + X_2 + X_3$
   (d) $X_1 + 2X_2 - X_3$
   (e) $3X_1 - 4X_2$ si X_1 y X_2 son variables aleatorias independientes.

10. Con el vector aleatorio $X' = [X_1, X_2, X_3, X_4]$ con vector de media $\mu_x' = [3, 2, -2, 0]$ y matriz de varianza-covarianza

$$\mathbf{\Sigma_X} = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{bmatrix}$$

$$\text{Sea } \mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & -2 & 0 \\ 1 & 1 & 1 & -3 \end{bmatrix}$$

(a) Encuentre $E(AX)$, la media de AX.

(b) Determine Cov (AX), las varianzas y covarianzas de AX.

(c) ¿Cuáles pares de combinaciones lineales tienen covarianza cero?