MATH 250: Mathematical Data Visualization
Spring 2025

Homework 5: Inference and Multivariate Statistics

For all questions below, you should show all work needed to reach your answer. You may collaborate with your classmates and consult external resources, but you should write and submit your own answer. **Any classmates with whom you collaborate should be credited at the top of your submission. Similarly, if you consult any external references, you should cite them clearly and explicitly.**

1. Suppose X and Y are independent $N(0, 1)$ random variables. Derive the distribution of the random vector $\begin{bmatrix} X+Y \\ X-Y \end{bmatrix}$.

2. Suppose $X_1, \dots, X_n, X_{n+1}$ are identically and independently distributed $N(\mu, 1)$ random variables.
   (a) Derive the conditional distribution of $X_{n+1} \mid X_1, \dots X_n$.
   (b) Suppose $X_1, \dots, X_n$ are unobserved, but the sample mean $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ is observed. Derive the conditional distribution of $X_{n+1} \mid \overline{X}_n$.
   (c) Suppose instead that for all $i \neq j$, the correlation between X_i and X_j is ρ for some $0 < \rho < 1$. Derive the conditional distribution of $X_{n+1} \mid \overline{X}_n$.

3. (Sherman-Morrison-Woodbury) Show that

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U\left(C^{-1} + VA^{-1}U\right)^{-1} VA^{-1}$$

   where A is an invertible $n \times n$ matrix , C is an invertible $k \times k$ matrix, U is $n \times k$ and V is $k \times n$.

4. (Inverse of a rank-1 update) Use the above identity to show that if A is an invertible $n \times n$ matrix and $\mathbf{u}$ and $\mathbf{v}$ are length n vectors:

$$\left(A + \mathbf{u}\mathbf{v}^{\top}\right)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{u}\mathbf{v}^{\top}A^{-1}}{1 + \mathbf{v}^{\top}A^{-1}\mathbf{u}}$$

5. In your own words, explain why using factor analysis to draw causal conclusions may be incorrect. Give an example to illustrate the issue.

6. Complete the **Factor Analysis R Lab** and submit via Gradescope.