# Statistical Machine Learning - HW 1

## Problem 1

Let $\mathbf{X} = (X_1, X_2, X_3)^T$ be distributed as $\mathbf{X} \sim MVN(\mu, \Sigma)$, where

$$\mu = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

Which of the following pairs of random variables are independent? Explain why.

a. $X_1$ and $X_2$
b. $X_1$ and $X_3$
c. $X_2$ and $X_3$
d. $(X_1, X_3)$ and $X_2$
e. $X_1$ and $X_1 + 3X_2 - 2X_3$

## Problem 2

Suppose $\mathbf{X}$ is a random vector with mean $\mu$. Prove that

$$E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^T] = E(\mathbf{X}\mathbf{X}^T) - \mu\mu^T.$$

## Problem 3

Let $\mathbf{Z} \in \mathbb{R}^p$ be distributed as $\mathbf{Z} \sim MVN(\mathbf{0}, I_p)$. Given $\mu \in \mathbb{R}^p$ and a positive definite $p \times p$ matrix $\Sigma$, derive methods to transform $\mathbf{Z}$ into a $p$-variate normal random vector with mean $\mu$ and covariance $\Sigma$ using:

a. The spectral decomposition, $\Sigma = V\Lambda V^T$, where $V^T V = I_p$ and $\Lambda$ is diagonal.
b. The Cholesky decomposition, $\Sigma = LL^T$, where $L$ is lower-triangular.

## Problem 4

Write Python (or R) code that uses the covariance matrix $\Sigma$ from **Problem 1** to generate **1000** trivariate normal random vectors using both methods from **Problem 3(a)** and **3(b)**. Compare the associated **sample covariance matrices** from the two methods.

## Problem 5

**Exercise #7** from Chapter **2** of the textbook: "7. The table below provides a training data set containing six …"