# Recitation 12. December 3

***Focus: Probability, Statistics and Markov Chains***
Consider running an experiment $n$ times measuring a certain quantity. We call these values $x_1, x_2 \dots x_n$ ***samples***. The collection of these is known as ***data set***.
The ***mean*** of a data set is given by $\mu = \frac{1}{n}(x_1 + \cdots + x_n)$. The ***variance*** of the data set is $\Sigma = \frac{1}{n-1}((x_1 - \mu)^2 + \cdots + (x_n - \mu)^2)$. Similarly for a second data set $y_1, \dots y_n$ the covariance between $x$'s and $y$'s is given by $\Sigma_{xy} = \frac{1}{n-1}((x_1 - \mu)(y_1 - \nu) + \cdots + (x_n - \mu)(y_n - \nu))$, where $\nu$ is the average of the $y$'s.

Let $P = I - \mathbf{o}\mathbf{o}^T/\mathbf{o}^T\mathbf{o}$, where $\mathbf{o} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, ie the projection matrix to the orthogonal complement of $\mathbf{o}$. Let $A = \begin{bmatrix} x_1 & y_1 & z_1 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & z_n & \dots \end{bmatrix}$ a matrix of different data sets. Then the covariance matrix is computed by $K = \frac{A^T P A}{n-1}$.
A ***Markov matrix*** $M$ is a square matrix with non-negative entries whose columns add up to 1. It models a ***Markov process*** which has $n$ states and after each step the $i$th column of $M$ gives the probability of moving from the $i$th state to the other. So if we start with a distribution $v$ computing $Mv$ gives the probability distribution after one step. Using eigenvalues and eigenvectors we can compute the steady state given by the limit $M^k v$ as $k \to \infty$

1. Consider the following measurements of temperature and pressure (measured in some units) given by $Temp = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ and $Press = \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix}$
   - Compute the covariance matrix.
   - Find linear combinations of temperature and pressure that are independent

   **Solution:**

2. Consider three independent random variable $X_1$, $X_2$ and $X_3$. Assume they have variances $\Sigma_1$, $\Sigma_2$ and $\Sigma_3$ and means $\mu_1$, $\mu_2$ and $\mu_3$.
   - What is the covariance matrix of these variables?
   - What is the variance of $X_1 + X_2 + X_3$?
   - What is the covariance of $X_1$ and $X_1 + X_2 + X_3$?

   **Solution:**

3. Consider the matrix

$$\begin{bmatrix} 1 & 2 \\ a & b \end{bmatrix}$$

For some constants $a$ and $b$ Suppose it is a covariance matrix of two random variables $X$ and $Y$.

- What can you say about $a$ and $b$?
- What can you say if on top of the above you know that there is some linear combination of $X$ and $Y$ that are constant?

**Solution:**

4. Consider the Markov process with two states up and down.

- If the state is up the next step is equaly likely to be up or down.
- If the state is down the next step is up.

1. Find the Markov matrix describing this Markov process.
2. If we start in the up state, what is the probability distribution after $k$ steps?
3. What is the steady state?

**Solution:**