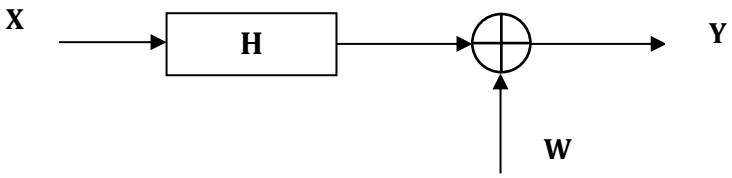
**ELEC 642 Machine Learning for Signal Processing**
**Homework #2**
**Due: 07.04.2020**

## Supervised Signal Restoration

1. Assume that we observe the signal $\mathbf{Y} = \mathbf{HX} + \mathbf{W}$ over an LTI causal channel given in the following figure. Here $\mathbf{H}$ is a convolution matrix constituted by the system impulse response $h[n]$ with length 5. The noise $\mathbf{W} \in \mathbb{R}^{256}$ is a white Gaussian vector with zero mean and variance $\sigma_w^2$. $\mathbf{X}$ and $\mathbf{W}$ are assumed to be uncorrelated.




We assume that the signal $\mathbf{X}$ is a Gaussian random vector with mean $\boldsymbol{\mu}_x$ and covariance matrix $\boldsymbol{\Sigma}_x$. $M = 254$ input-output pairs $\{\mathbf{x}_i, \mathbf{y}_i\}$ are given in the attached mat files. In the mat file, the $i$th column of matrix `x` and `y` correspond to $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively. Apply 10-fold cross-validation scheme to prepare the training and test samples and do the followings:

   a. Learn the parameters $\boldsymbol{\theta} = \{\mathbf{H}, \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \sigma_w^2\}$ of the observation model given $\{\mathbf{x}_i, \mathbf{y}_i\}$ pairs for $i = 1,2,\dots,N_A$ where $N_A$ is the number of training samples. To construct convolution matrix, you can use `toeplitz` command in MATLAB. Since the number of samples are not enough to estimate the covariance matrix, use the following MAP estimator

$$\widehat{\boldsymbol{\Sigma}}_x = 10^{-4}\mathbf{I} + \frac{1}{N_A}\sum_{i=1}^{N_A}(\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_x)(\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_x)^T$$

   b. Given test samples $\mathbf{y}_j, j = 1,2,\dots,N_B$, find the MAP estimate of the signal $\mathbf{x}_j$ using the estimated parameters set $\widehat{\boldsymbol{\theta}}$. Calculate the average peak signal-to-noise ratio (PSNR) such that

$$PSNR = 20\log_{10}\frac{\max(\mathbf{x})}{\sqrt{MSE}}$$

   where

$$MSE = \frac{1}{256N_B}\sum_{j=1}^{N_B}\left\|\mathbf{x}_j - \hat{\mathbf{x}}_j\right\|^2$$

   c. Repeat the process in (a) and (b) 10 times. Each time select a different part of the data for testing and the remaining 9 parts for training. List the best, worst and mean PSNR value in a table.

   d. Plot one of the estimates $\hat{\mathbf{x}}_j$ along with the corresponding ground-truth $\mathbf{x}_j$ and the observation $\mathbf{y}_j$.