Intro to Signal Processing – HW4

Yinon Goldshtein , Roee Idan

Submitted to January 26 2023

Part I
Theory

Question 1 – Inverting the Second Derivative Operator

a

$$H = \begin{pmatrix} \frac{-5}{2} & \frac{4}{3} & \frac{-1}{12} & 0 & \cdots & 0 & \frac{-1}{12} & \frac{4}{3} \\ \frac{4}{3} & \frac{-5}{2} & \frac{4}{3} & \frac{-1}{12} & 0 & \cdots & 0 & \frac{-1}{12} \\ \frac{-1}{12} & \frac{4}{3} & \frac{-5}{2} & \frac{4}{3} & \ddots & 0 & \cdots & 0 \\ 0 & \frac{-1}{12} & \frac{4}{3} & \frac{-5}{2} & \frac{4}{3} & \cdots & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \frac{4}{3} & \frac{-1}{12} & 0 & \cdots & 0 & \frac{-1}{12} & \frac{4}{3} & \frac{-5}{2} \end{pmatrix} \tag{1}$$

b

Because H is a circulant we get that it is diagnalizable with the DFT matrix and so $H = DFT \times \Lambda \times DFT^*$. We know that the first row of H_0 is

$$H_0 = \begin{pmatrix} \frac{-5}{2} \\ \frac{4}{3} \\ \frac{-1}{12} \\ 0 \\ \vdots \\ 0 \\ \frac{-1}{12} \\ \frac{4}{3} \end{pmatrix} \tag{2}$$

From here we can deduce that

$$\Lambda = \sqrt{M} \times DFT^* \times H_0 = \begin{pmatrix} \sum_{l=0}^{M-1} W_M^{0,l} H_0 \\ \sum_{l=0}^{M-1} W_M^{1,l} H_0 \\ \vdots \\ \sum_{l=0}^{M-1} W_M^{M-1,l} H_0 \end{pmatrix} = \begin{pmatrix} 0 \\ \sum_{l=0}^{M-1} W_M^{1,l} H_0 \\ \vdots \\ \sum_{l=0}^{M-1} W_M^{M-1,l} H_0 \end{pmatrix} \tag{3}$$

We know that the rank of H is $M-1$ and so we can deduce that Λ_0 is the only zero and beacause we know that H is symmetric we can deduce that all the eigenvalues are real and so the Psuedo inverse of H will be

$$H^+ = DFT \times \Lambda^+ \times DFT^* \tag{4}$$

where we define Λ^+ to be 0 in all places that aren't the diagonal and in the diagonal:

$$diag(\Lambda^+) = \begin{pmatrix} 0 \\ \frac{1}{\sum_{l=0}^{M-1} W_M^{1,l} H_0} \\ \vdots \\ \frac{1}{\sum_{l=0}^{M-1} W_M^{M-1,l} H_0} \end{pmatrix} \tag{5}$$

c

Using the previously provided inverse filter φ cannot be perfectly recovered due to the fact that H has a null space which from what we learnt is not invertible and from here we can deduce that any signal that part of its component will be in the null space cannot be recovered. So for example we can build a signal which is a scaled version of the eigenvector that is corresponding to the eigenvalue of 0 and which will be the first row of DFT^* and so for $c \neq 0$ we get that

$$\varphi := c \begin{pmatrix} W_M^0 \\ W_M^0 \\ \vdots \\ W_M^0 \end{pmatrix} = \begin{pmatrix} c \\ c \\ \vdots \\ c \end{pmatrix} \tag{6}$$

We can see that for this φ we get that $H \times \varphi = 0$ and also for all $c \neq 0$ we will get that

$$\varphi \neq H^+ \times H \times \varphi = 0 \tag{7}$$

Question 2 – Let's Randomise

a

Let us look at the expected value for each value and let us show that they are all 0. Let us look first $i \leq \frac{N}{2}$ we get that :

$$E[\varphi_i] = E[\varphi_i|K=i]P(K=i) + E[\varphi_i|K\neq i]P(K\neq i) = E[M+L_1]P(K=i) + E[M]P(K\neq i) =$$
$$= (E[M]+E[L_1])P(K=i) + E[M]P(K\neq i) = 0 \tag{8}$$

Let us look first $i > \frac{N}{2}$ we get that :

$$E[\varphi_i] = E[\varphi_i|K=i]P(K=i) + E[\varphi_i|K\neq i]P(K\neq i) = E[M+L_2]P(K=i) + E[M]P(K\neq i) =$$
$$= (E[M]+E[L_2])P(K=i) + E[M]P(K\neq i) = 0 \tag{9}$$

And from here can deduce that random vector φ has a zero mean

b

From the question's information we get that R_φ matrix will $R_\varphi = E[\varphi\varphi^T]$ let us denote in R_φ $r_{i.j}$ as the element in the place i,j. Let us note that M,L are independent and also that $E[M]=E[L]=0$. When $i\neq j$ we get that:

$$r_{i,j} = E[\varphi_i\varphi_j] = E[\varphi_i\varphi_j|K=i]P(K=i) + E[\varphi_i\varphi_j|K=j]P(K=j) + E[\varphi_i\varphi_j|K\neq i,j]P(K\neq i,j) =$$
$$= E[(M+L)M]P(K=i) + E[M(M+L)]P(K=j) + E[M^2]P(K\neq i,j) =$$
$$= (E[M^2]+E[L]E[M])P(K=i) + (E[M^2]+E[M]E[L])P(K=j) + E[M^2]P(K\neq i,j) =$$
$$= E[M^2]P(K=i) + P(K=j) + P(K\neq i,j) = E[M^2] = c \tag{10}$$

When $i=j$ we get that:

$$r_{i,j} = E[\varphi_i\varphi_i|K=i]P(K=i) + E[\varphi_i\varphi_i|K\neq i]P(K\neq i) = E[(M+L)(M+L)]P(K=i) + E[M^2]P(K\neq i) =$$
$$= (E[M^2]+2E[M]E[L]+E[L^2])P(K=i) + E[M^2]P(K\neq i) \tag{11}$$

And from here we get that when $i \leq \frac{N}{2}$

$$r_{i,j} = E[M^2] + \frac{E[L_1^2]}{N} = a + c \tag{12}$$

and when $i > \frac{N}{2}$

$$r_{i,j} = E[M^2] + \frac{E[L_2^2]}{N} = b + c \tag{13}$$

c

In order for the PCA to be DFT^* we need that R_φ should be diagonalizable with DFT^* which will occur if and only if R_φ is circulant and so for that to happened we will need that $a=b$

Question 3 – Let's Randomise Again!

Now the signal φ is:

$$\varphi = \begin{bmatrix} M & \cdots & M+L & M & \cdots & M & M+L & M & \cdots & M \end{bmatrix}^T$$

Where the $M+L$ terms appear in the K-th and the $(K+\frac{N}{2})$-th elements.

a

Available random variables: $K \sim U(\{1,\dots,\frac{N}{2}\}), \mathbb{E}(M)=0, \mathbb{E}(M^2)=c, \mathbb{E}(L)=0, \mathbb{E}(L^2)=\frac{N}{2}(1-c)$.
c is a real constant $c \in (0,1)$.

We know that K, M, L are independent, therefore:

$$\begin{aligned}
&\mathbb{E}(L\cdot M) = \mathbb{E}(L)\cdot\mathbb{E}(M) = 0\\
&\mathbb{E}(M(M+L)) = \mathbb{E}(M^2)+\mathbb{E}(LM) = \mathbb{E}(M^2) = c\\
&\mathbb{E}((M+L)^2) = \mathbb{E}(M^2+2LM+L^2) = \mathbb{E}(M^2)+\mathbb{E}(2LM)+\mathbb{E}(L^2) = c+\frac{N}{2}(1-c)
\end{aligned}$$

Now we can calculate the autocorrelation matrix by: $R_\varphi = \mathbb{E}(\varphi\varphi^T)$

$$\varphi\varphi^T = \begin{pmatrix}
M^2 & \cdots & M^2 & M(M+L) & M^2 & \cdots & M^2 & M(M+L) & M^2 & \cdots M^2\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
M(M+L) & \cdots & \cdots & (M+L)^2 & \cdots & \cdots & M(M+L) & (M+L)^2 & \cdots & M(M+L)\\
M^2 & \cdots & & & & & & & \cdots & M^2
\end{pmatrix}$$

So actually, the K and $K+\frac{N}{2}$ columns and rows are:

$$[\varphi\varphi^T]_{i,j} = \begin{cases} (M+L)^2, & i\in\{K, K+\frac{N}{2}\}, j\in\{K,K+\frac{N}{2}\}\\ M(M+L), & i,j\in\{K,K+\frac{N}{2}\}\\ M^2, & otherwise\end{cases}$$

And so the auto-correlation matrix, which is $R_\varphi = E[\varphi\varphi^T]$ is:

$$\begin{aligned}
&[R_\varphi]_{i,j} = \mathbb{E}(\varphi_i\cdot\varphi_j)\\
&(1)\quad i=j \implies R_{i,j} = \mathbb{E}(\varphi_i^2) =\\
&\mathbb{E}(\varphi_i^2 \mid i\neq K, K+\frac{N}{2})\cdot P(i\neq K,K+\frac{N}{2}) + \mathbb{E}(\varphi_i^2 \mid i=K,K+\frac{N}{2})\cdot P(i=K,K+\frac{N}{2}) =\\
&E(M^2)\cdot\frac{N-2}{N} + E((M+L)^2)\cdot\frac{2}{N} = c\cdot\frac{N-2}{N}+\frac{2}{N}\cdot(c+\frac{N}{2}\cdot(1-c)) =\\
&c+1-c = 1\\
&(2)\quad i-j = \frac{N}{2} \mod N (\text{The N/2 diagonal}) \implies R_{i,j} = \mathbb{E}(\varphi_i\varphi_j) =\\
&\mathbb{E}(\varphi_i\varphi_j \mid i,j\neq K,K+\frac{N}{2})\cdot P(i,j\neq K,K+\frac{N}{2}) + \mathbb{E}(\varphi_i\varphi_j \mid i,j=K,K+\frac{N}{2})\cdot P(i,j=K,K+\frac{N}{2}) \underbrace{=}_{*}\\
&\mathbb{E}(M^2)\cdot\frac{N-2}{N}+\mathbb{E}((M+L)^2)\cdot\frac{2}{N} = c\cdot\frac{N-2}{N}+\frac{2}{N}\cdot(c+\frac{N}{2}\cdot(1-c)) = 1\\
&(3)\quad i-j\neq\frac{N}{2} \mod N \wedge i\neq j (\text{otherwise}) \implies R_{i,j} = \mathbb{E}(\varphi_i\varphi_j) =\\
&\mathbb{E}(\varphi_i\varphi_j \mid i=K,K+\frac{N}{2})\cdot P(i=K,K+\frac{N}{2}) + \mathbb{E}(\varphi_i\varphi_j \mid j=K,K+\frac{N}{2})\cdot P(j=K,K+\frac{N}{2}) +\\
&\mathbb{E}(\varphi_i\varphi_j \mid i,j,K+\frac{N}{2})\cdot P(i,j\neq K,K+\frac{N}{2}) =
\end{aligned}$$

$$\mathbb{E}(M(L+M)) \cdot \frac{2}{N} \cdot 2 + \frac{N-4}{N} \cdot \mathbb{E}(M^2) = c \cdot \frac{4}{N} + c \cdot \frac{N-4}{N} = c$$

So in conclusion:

$$[R_\varphi]_{i,j} = \begin{cases} 1, & \text{Main diagonal, or } \pm\frac{N}{2} \text{ diagonals} \\ c, & otherwise \end{cases}$$

Row 0 of R_φ is:

$$\begin{pmatrix} 1 & c & \cdots & c & 1 & c & \cdots & c \end{pmatrix}$$

And every following row is the previous row, shifted right by one element (in row 2 the 1 is in the second position).

Therefore, R_φ is circulant with a period of $\frac{N}{2}$.

b

We saw in HW3 that the eigenvalues of R_φ, stacked in a column, equals to the product of the DFT^* matrix and the first row of R_φ rewritten in column form, normalized by the root of the size of R_φ. Therefore:

$$\begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_{N-1} \end{pmatrix} = \sqrt{(N)} \cdot DFT^* \cdot \begin{pmatrix} 1 & c & \cdots & c & 1 & c & \cdots & c \end{pmatrix}^T$$

Question 4 – Wiener Filter

a

We know that $\varphi^* = H\varphi + n$ and so the autocorrelation will be defined by:

$$\begin{aligned} R_{\varphi^*} &= E(\varphi\varphi^*) = E((H\varphi+n)(H\varphi+n)^*) = E((H\varphi+n)(H^*\varphi^*+n^*)) = \\ &= E(H\varphi\varphi^*H^* + H\varphi n^* + n\varphi^*H^* + nn^*) = \\ &= E(H\varphi\varphi^*H^*) + E(H\varphi n^*) + E(n\varphi^*H^*) + E(nn^*) = \\ &= HE(\varphi)E(\varphi^*)H^* + HE(\varphi)E(n^*) + E(n)E(\varphi^*)H^* + E(nn^*) = \\ &= HR_\varphi H^* + R_n \end{aligned} \tag{14}$$

b

For the case $R_n = \sigma_n^2$I is from what we learned in class to be and from what we saw in the previous subsection we get that

$$W = R_\varphi H^*(HR_\varphi H^* + \sigma_n^2 I)^{-1} == R_\varphi H^* R_{\varphi^*}^{-1} \tag{15}$$

c

We get that for $A = DFT\Lambda DFT^*$ because by definition $W_N^l = W_N^{l(modN)}$

$$a_{i,j} = \sum_{k=0}^{N-1} W_N^{-ik}\lambda_k W_N^{-jk} = \sum_{k=0}^{N-1} \lambda_k W_N^{k(j-i)} = \sum_{k=0}^{N-1} \lambda_k W_N^{k(j-i)(modN)} = a_{0,(j-i)(modN)} \tag{16}$$

and from here we can deduce that A will be a circulant matrix.

d

The Wiener filter is not a shift invariant system, for example for the case that $R_\varphi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $H_\varphi = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\sigma_n^2 = 1$ we will get from calculation that the filter will be from the equation

$$W = R_\varphi H^*(HR_\varphi H^* + \sigma_n^2 I)^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})^{-1} = \begin{pmatrix} 0 & 0 \\ 0.5 & 0 \end{pmatrix} \tag{17}$$

Note that the Wiener filter is not circulant and so from what we learned we can deduce that it is not a shift invariant.

e

Let us note A, B as circulant matrixes. we get that also A^*, B^* are also circulant and we get as well that

$$A^{-1} = ((DFT)\Lambda(DFT^*))^{-1} = (DFT)\Lambda^{-1}(DFT^*) \tag{18}$$

and so we get that A^{-1}, B^{-1} are also circunt also we can calculate that

$$A + B = (DFT)\Lambda_A(DFT^*) + (DFT)\Lambda_B(DFT^*) = (DFT)(\Lambda_A + \Lambda_B)(DFT^*) \tag{19}$$

and so we get that $A + B$ is also circulant also we can calculate that

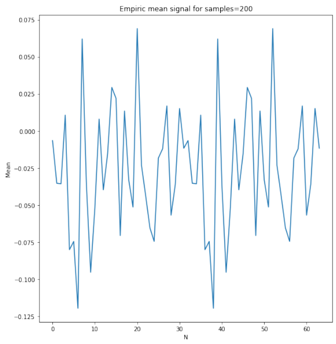
$$AB = (DFT)\Lambda_A(DFT^*) \times (DFT)\Lambda_B(DFT^*) = (DFT)(\Lambda_A\Lambda_B)(DFT^*) \tag{20}$$

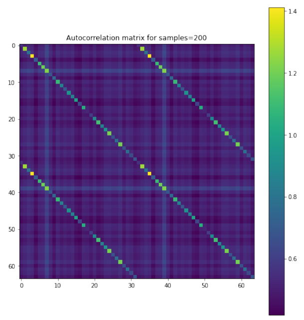
and so we get that AB is also circulant. also not that for every σ_n^2 we get that $\sigma_n^2 I$ is circulant. From here we can deduce that the condition will be that R_φ and H will be circulant inorder for the Wiener filter to be shift-invariant.

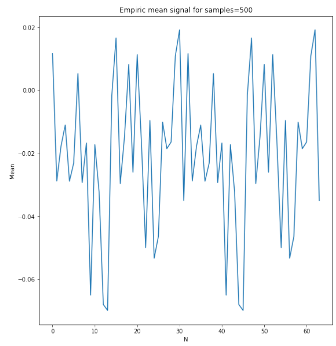
Part II
Implementation

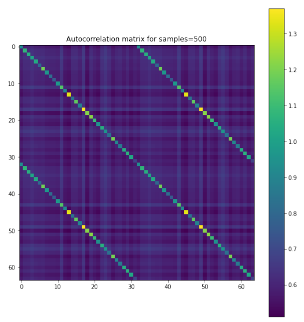
a

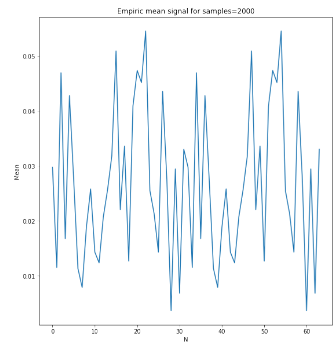
Given the possible sample sizes of $[200, 500, 2000, 5000]$, the mean signal and autocorrelation matrices are:

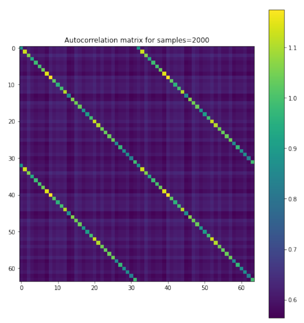


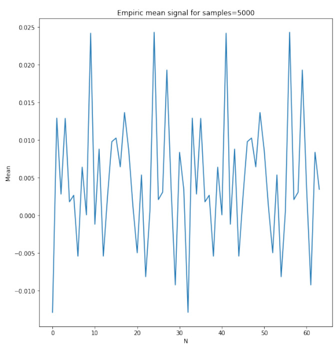


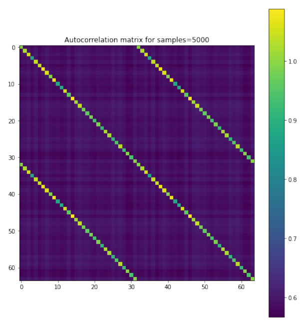












As expected, we can see that our theory holds more when the sample size increases.
Looking at our max sample size which is $num = 5000$, we can see in the auto-correlation matrix that

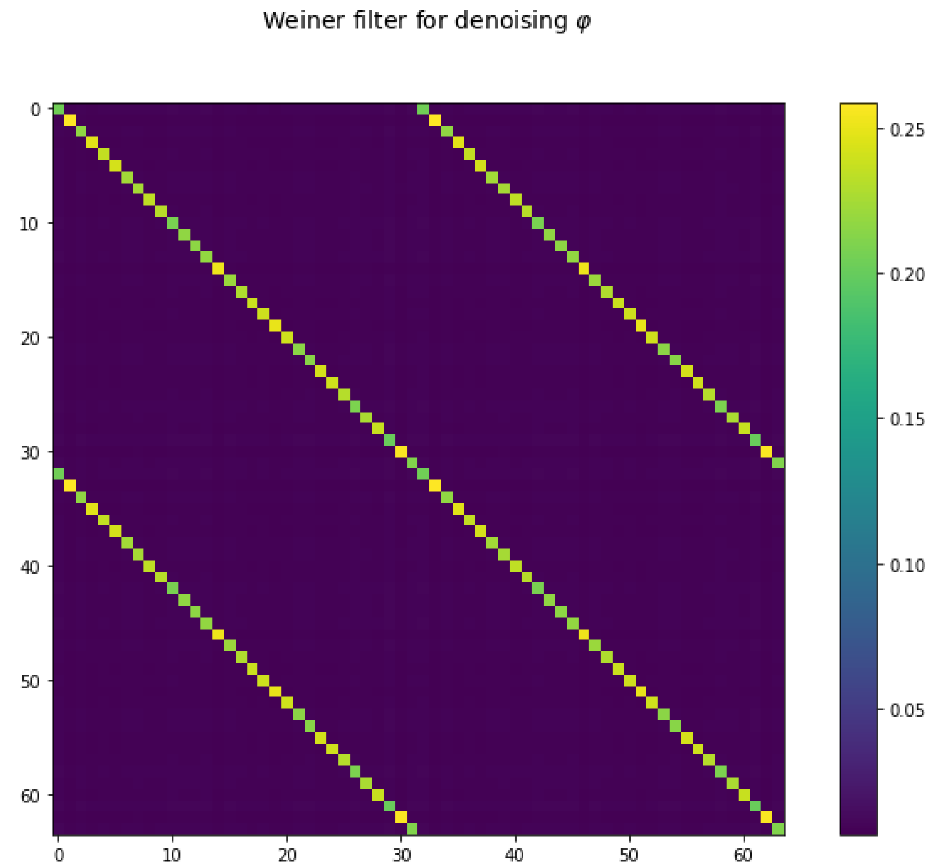
almost all values not on the three "special" diagonals are approximately $\approx c = 0.6$, and the values on the diagonals are approximately ≈ 1.

In our mean signal figure, we can see that the mean is 0 ± 0.025, which is centered quite good.

From here, we proceed with our values using sample size=5000.

b

The Wiener filter matrix that we obtained:



In this section, our linear degradation matrix is $H = I$, so the Weiner filter's formal calculation will be: $W = R_{\varphi} \cdot \left(R_{\varphi} + \sigma_n^2 \cdot I\right)^{-1}$

We can also see that the Weiner matrix is circulant, in according to Q4 in the dry part - where if the auto-correlation is circulant, the weiner matrix is circulant.

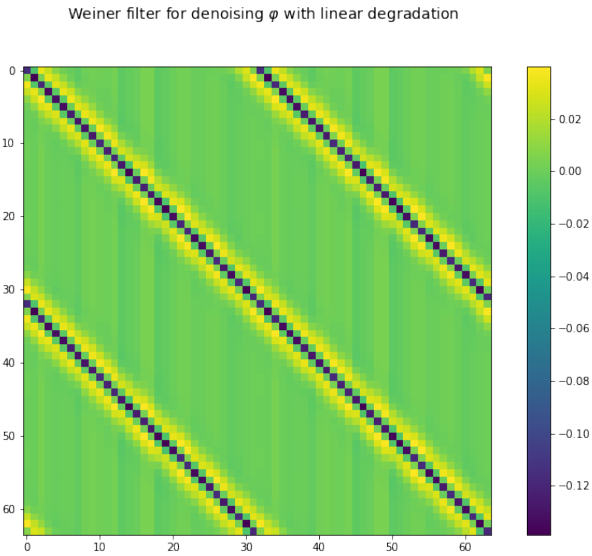
This is a plot of 10 random signal samples, with their corresponding noisy and denoised signals:

Clean, noisy and denoised signals (section b)

The mean MSE of all Weiner-denoised signals with respect to the original signals is: 0.229
This result is good, as the MSE is an order of magnitude lower than the variance of the noise.

c

The Wiener filter matrix that we obtained:



In this section, our linear degradation matrix is a circulant matrix with a first row of $\begin{bmatrix}-\frac{5}{2} & \frac{4}{3} & -\frac{1}{12} & 0 & \ldots & 0 & -\frac{1}{12} & \frac{4}{3}\end{bmatrix}$, so the Weiner filter's formal calculation will be: $W = R_\varphi \cdot H^T \cdot \left(H \cdot R_\varphi \cdot H^T + \sigma_n^2 \cdot I\right)^{-1}$

We can also see that the Weiner matrix is circulant, in according to Q4 in the dry part - where if the auto-correlation is circulant, the weiner matrix is circulant.

This is a plot of 10 random signal samples, with their corresponding noisy and denoised signals:

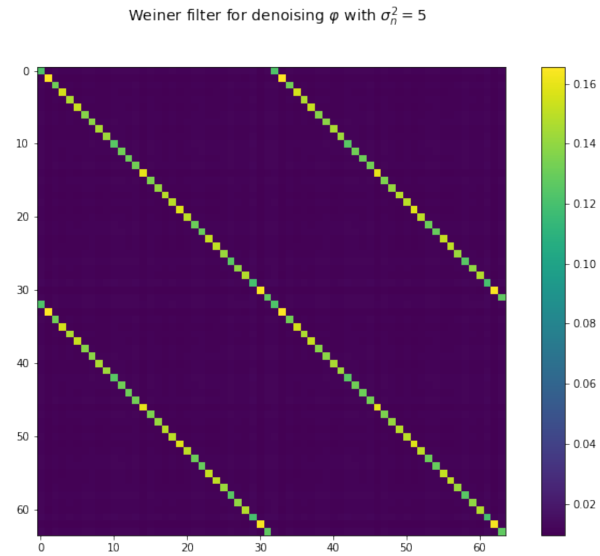
The mean MSE of all Weiner-denoised signals with respect to the original signals is: 1.275

This result is worse then before, when we calculated the MSE without the added linear degradation. This is because H is not invertible, because now the parts of the vectors that corresponds to the eigenvector of the eigenvalue 0 are degraded and lost.

d

Without linear degradation

The Wiener filter matrix that we obtained:



In this section, our linear degradation matrix is $H = I$, so the Weiner filter's formal calculation will be: $W = R_\varphi \cdot \left(R_\varphi + \sigma_n^2 \cdot I\right)^{-1}$

We can also see that the Weiner matrix is circulant, in according to Q4 in the dry part - where if the auto-correlation is circulant, the weiner matrix is circulant.

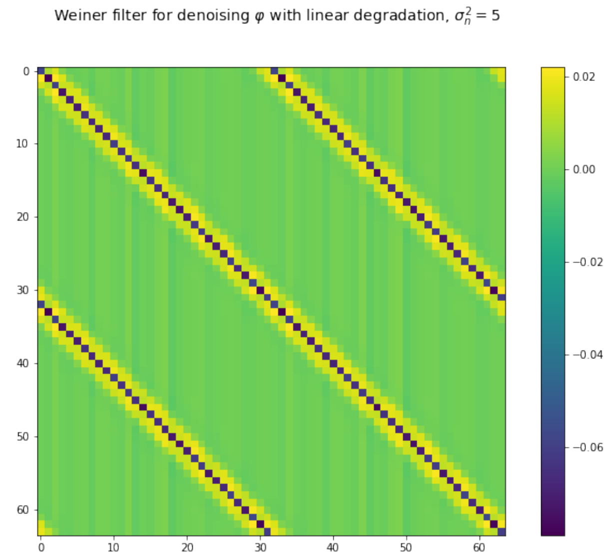
This is a plot of 10 random signal samples, with their corresponding noisy and denoised signals:

Clean, noisy and denoised signals (section d1)

The mean MSE of all Weiner-denoised signals with respect to the original signals is: 0.447
This is higher than the section b MSE, and can be explained by the higher variance.

With linear degradation

The Wiener filter matrix that we obtained:



In this section, our linear degradation matrix is a circulant matrix with a first row of $\begin{bmatrix} -\frac{5}{2} & \frac{4}{3} & -\frac{1}{12} & 0 & \dots & 0 & -\frac{1}{12} & \frac{4}{3} \end{bmatrix}$, so the Weiner filter's formal calculation will be:
$W = R_\varphi \cdot H^T \cdot \left(H \cdot R_\varphi \cdot H^T + \sigma_n^2 \cdot I\right)^{-1}$

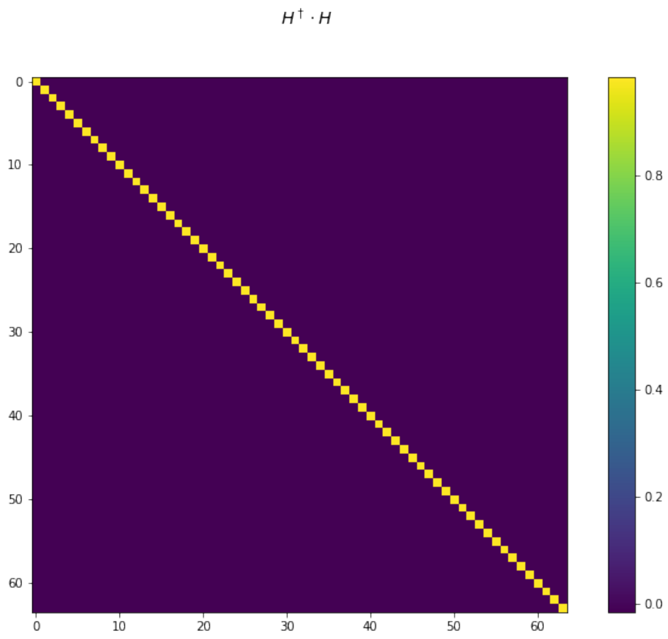
We can also see that the Weiner matrix is circulant, in according to Q4 in the dry part - where if the auto-correlation is circulant, the weiner matrix is circulant.

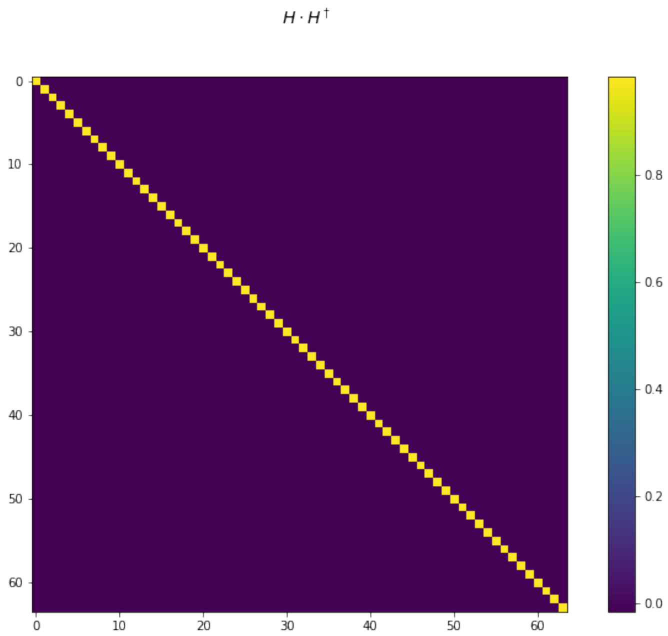
This is a plot of 10 random signal samples, with their corresponding noisy and denoised signals:

The mean MSE of all Weiner-denoised signals with respect to the original signals is: 1.173
This is higher than the section c MSE, and can be explained by the higher variance.

e

We computed the pseudo-inverse of H with: *np.linalg.pinv*, which is the moore-penrose inverse.





We can clearly see that $H^\dagger H = HH^\dagger = I$.

For our signals, given the signals:

$$\begin{aligned}
&\psi_1 = [100, 0, \dots, 0] \\
&\psi_2 = [1, 1, \dots, 1] = ones(N = 64) \\
&\phi_1 = 10 \cdot ones(N) + \psi_1 \quad \text{\# first signal} \\
&\phi_2 = 100 \cdot ones(N) + \psi_1 \quad \text{\# second signal}
\end{aligned}$$

We know from Q1 that the first column if the DFT^* matrix is the eigenvector of H that corresponds to the first eigenvalue - which is 0, and the first column of the DFT^* is $ones(N)$

Therefore, we get that $\lambda = 0$ is an eigenvalue of $H^\dagger$ with the eigenvector $ones(N)$ because $H^\dagger$ is circulant and diagonalizable by DFT^*.

Because of that, we get that $10 \cdot ones(N)$ and $100 \cdot ones(N)$ are eigenvectors of $H^\dagger$ with eigenvalue 0.

So:

$$\begin{aligned}
&H^\dagger \cdot \phi_1 = H^\dagger \cdot (10ones(N) + \psi_1) = \underbrace{H^\dagger \cdot 10 \cdot ones(N)}_{=0} + H^\dagger \cdot \psi_1 = H^\dagger \cdot \psi_1 \\
&H^\dagger \cdot \phi_2 = H^\dagger \cdot (100ones(N) + \psi_1) = \underbrace{H^\dagger \cdot 100 \cdot ones(N)}_{=0} + H^\dagger \cdot \psi_1 = H^\dagger \cdot \psi_1 \\
&\implies H^\dagger \cdot \phi_1 = H^\dagger \cdot \phi_2 \\
&Also: \quad \|\phi_1 - \phi_2\|_2 = 90^2 \cdot 64 \geq 256
\end{aligned}$$

We can compute the norm $\|\phi_1 - \phi_2\|_2 = 720$

And the $L2$ norm of $\|H^\dagger \phi_1 - H^\dagger \phi_2\|_2 = 9.9788 \cdot 10^{-9}$, as expected.

The plots we got for the signals: