Signal Processing, Representation, Modeling, and Analysis

Alfred M. Bruckstein & Ron Kimmel

Geometric Image Processing Lab.
CS Dept. Technion

Winter 2017-2018
Chapter 12

Optimal Operators (Part I)
Next, we will learn to design linear operator that best recover signals from their linearly distorted versions. So far we dealt with best bases that minimize the representation error in partial approximations using the first $K$ elements of the given basis instead of the basis as a whole, with $N$ components in the vector case and $\infty$ in the continuous one.

It is useful for data compression, where the principal component analysis, realized by the Korhunen-Loeve transform, is often utilized for the design of efficient storage, information communication, to name just a few of potential applications that could benefit from the PCA.

Another important goal of signal processing is signal recovery after systematic distortions and contamination by noise. Noise can be thought of as an adversary signal from a class different from that of the clean signal. Let us limit ourselves to distortions that can be represented by linear operators and additive noise which is statistically independent of the clean signal.
Signal, distortion and noise

\[
\begin{align*}
\{\psi_\omega\} & \rightarrow \psi_0 \rightarrow \mathcal{H}(\psi_0) \rightarrow \mathcal{H}(\psi_0) \\
\{n_\tilde{\omega}\} & \rightarrow n_{\tilde{\omega}_0} \\
\psi_{\text{Data}} & = \mathcal{H}(\psi_0) + n_{\tilde{\omega}_0}.
\end{align*}
\]

Given the signal \(\psi_{\text{Data}}^{\omega,\tilde{\omega}}\) we would like to design an operator \(\mathcal{M}(\cdot)\) that when applied to \(\psi_{\text{Data}}^{\omega,\tilde{\omega}}\) it provides \(\psi_0\).

Given \(\psi_{\omega,\tilde{\omega}}^{\text{Data}} = \mathcal{H}(\psi_\omega) + n_{\tilde{\omega}}\), find \(\mathcal{M}(\cdot)\) such that \(\hat{\psi}_{\omega,\tilde{\omega}} = \mathcal{M}(\psi_{\omega,\tilde{\omega}}^{\text{Data}})\) that minimizes \(\mathcal{E}^2\) defined as

\[
\Psi(\mathcal{E}^2) = \Psi(\|\psi_\omega - \hat{\psi}_{\omega,\tilde{\omega}}\|^2).
\]
We assume the signal is

- a realization of a random process with zero mean and given autocorrelation, namely,

\[
\begin{align*}
\{ \bar{\psi}_\omega \}_\omega &\in \Omega \\
\sim & \quad E_\omega(\bar{\psi}_\omega) = 0 \\
E_\omega(\bar{\psi}_\omega \bar{\psi}_\omega^*) &= \mathcal{R}_\psi,
\end{align*}
\]

- the distortion $H(\cdot)$ is a linear, and often shift invariant, operator $H_{LSI}$,

- the noise signal is a realization of a random process with zero mean and known autocorrelation $\sigma_n^2 I$,

\[
\begin{align*}
\{ \bar{n}_\omega \}_{\tilde{\omega}} &\in \tilde{\Omega} \\
\sim & \quad E_{\tilde{\omega}}(\bar{n}_{\tilde{\omega}}) = 0 \\
E_{\tilde{\omega}}(\bar{n}_{\tilde{\omega}} \bar{n}_{\tilde{\omega}}^*) &= \sigma_n^2 I,
\end{align*}
\]

- the noise and signal processes are independent,

- the noise is additive.
Assume $\mathcal{H}(\cdot)$ is given by a known matrix $\mathbb{H}$ or $h(y, \tau)$. Consider the following three cases.

- **Inverse filtering/deconvolution:** $\psi$ is our signal and $\psi^{Data}$ is its distorted version $\mathcal{H}(\psi)$.

- **Constrained deconvolution:** The energy of our signal $\psi$ is roughly known, and so does the second order statistics of the noise, and the distortion operator $\mathcal{H}(\cdot)$.

- **Optimal least squares estimation or Wiener Filtering:** $\psi$ is a realization of a random process with known second order statistics. $n_{\tilde{\omega}}$ is a realization of a noise random process and $\mathcal{H}(\cdot)$ is given. Here, we need to consider the stochastic properties of both $\psi_\omega$ and $n_{\tilde{\omega}}$.

We will design the best linear signal reconstruction operators for each case which are not best in the more general non-linear case.
Inverse Filtering and Pseudo-Inverse

Given $\psi^\text{Data} = \mathcal{H}(\psi)$ from which we would like to recover $\psi$. We would have loved to find $\mathcal{H}^{-1}$ so that

$$\mathcal{H}^{-1}(\mathcal{H}(\psi)) = \psi,$$

that is

$$\mathcal{H}^{-1}(\mathcal{H}(\cdot)) = \text{Identity}(\cdot).$$

Yet, distortions are rarely invertible.

Consider the continuous case $\psi(t) : [0, 1) \rightarrow [\psi_L, \psi_H]$ distorted by a linear operator described by $h(t, \tau)$

$$\psi^\text{Data}(t) = \int_0^1 \psi(\xi) h(t, \xi) d\xi.$$

Assume $h(t, \xi) = h_{SI}(t - \xi)$, that is $\mathcal{H}$ is a LSI operator. Then

$$\psi^\text{Data}(t) = \int_0^1 \psi(\xi) h(t - \xi) d\xi,$$

where we assume that both $\psi$ and $h$ can be periodically extended beyond $[0, 1)$. 
Then,

\[
\psi(t) = \sum_{k=-\infty}^{+\infty} \langle \psi(t), e^{i2\pi kt} \rangle \psi_k e^{i2\pi kt}
\]

\[
h(t) = \sum_{k=-\infty}^{+\infty} \langle h(t), e^{i2\pi kt} \rangle h_k e^{i2\pi kt},
\]

and thus,

\[
\psi^{\text{Data}}(t) = \int_0^1 \sum_k \sum_l \psi_k e^{i2\pi k \xi} h_l e^{i2\pi l(t-\xi)} d\xi
\]

\[
= \sum_k \sum_l \psi_k h_l \int_0^1 e^{i2\pi (k-l) \xi} d\xi \ e^{i2\pi lt} \delta_{k-l}
\]

\[
= \sum_k \sum_l \psi_k h_l \delta_{k-l} e^{i2\pi lt}
\]

\[
= \sum_k \psi_k h_k e^{i2\pi kt}.
\]
Therefore,

\[ \psi_{Data}(t) = \sum_{k=-\infty}^{+\infty} \psi_k e^{i2\pi kt} \]

where \( \psi_k = \psi_k h_k \).

By comparing coefficients \( \psi_{Data}^k = \psi_k h_k \). Given \( h_k \), we could recover \( \psi_k \) from \( \psi_{Data}^k \) by a linear time invariant operator \( M(\cdot) \) defined by \( m_k = \frac{1}{h_k} \). In principle, perfect recovery would be realized by

\[ M(\psi_{Data}^k) = \frac{1}{h_k} \psi_{Data}^k = \frac{1}{h_k} h_k \psi_k = \psi_k. \]

The problem is that for some \( k \) we could have \( h_k = 0 \), then, \( M_k(\cdot) \) expressed by \( m(t) \) can not be written as

\[ m(t) = \sum_{k=-\infty}^{+\infty} \frac{1}{h_k} e^{i2\pi kt}. \]
Dealing with zero coefficients

As $\psi_k^{Data} = \psi_k h_k$, we have that $\psi_k^{Data} = 0$ for all $k$'s where $h_k = 0$. These coefficients of the signal $\psi(t)$ can not be recovered. Therefore, when $\psi_k^{Data} = 0$ we may set $\hat{\psi}_k = 0$, in which case

$$m_k \equiv \begin{cases} \frac{1}{h_k} & \text{if } h_k \neq 0 \\ 0 & \text{if } h_k = 0, \end{cases}$$

that yields

$$\hat{\psi}_k = \begin{cases} \psi_k & \text{if } h_k \neq 0 \\ 0 & \text{if } h_k = 0. \end{cases}$$

For this $M(\cdot)$, the estimation error is

$$\|\psi(t) - \hat{\psi}(t)\|^2 = \frac{1}{1} \int_0^1 \left( \sum_{k=-\infty}^{+\infty} (\psi_k - \hat{\psi}_k) e^{i2\pi kt} \right)^2 dt$$

$$= \sum_{k=-\infty}^{+\infty} (\psi_k - \hat{\psi}_k)^2 \int_0^1 (e^{i2\pi kt})^2 dt = \sum_{k \mid h_k=0} \psi_k^2.$$
Without additional information about \( \psi(t) \) there is not much we could do. Note that the energy of \( \hat{\psi}(t) \) is

\[
\int_0^1 (\hat{\psi}(t))^2 \, dt = \int_0^1 \left( \sum_{k=-\infty}^{+\infty} \hat{\psi}_k e^{i2\pi kt} \right)^2 \, dt
\]

\[
= \sum_{k \mid h_k \neq 0} \psi_k^2 + \sum_{k \mid h_k = 0} 0^2.
\]

Any different than zero selection for \( m_k \) at places with \( h_k = 0 \) would increase the energy of \( \hat{\psi}(t) \). We thus obtained the minimal energy estimator. Formally, we solved

\[
\min \int_0^1 \hat{\psi}^2(t) \, dt
\]

subject to

\[
\int_0^1 \hat{\psi}(\xi) h(t - \xi) \, d\xi = \psi^{Data}(t).
\]
The discrete inverse filtering problem

Given $\bar{\psi}$, we get $\psi^{Data} = \mathbb{H}\bar{\psi}$ from which we want to recover $\bar{\psi}$. When $\mathbb{H}_{N \times N}$ is invertible, we can have

$$\bar{\psi} = \mathbb{H}^{-1}\psi^{Data}.$$ 

If $\mathbb{H}$ is not invertible, we rely on the *singular value decomposition* (SVD) of $\mathbb{H}$. 
Any matrix $M_{m \times n}$ in $\mathbb{C}^{m \times n}$ or $\mathbb{R}^{m \times n}$ has a singular value decomposition

$$M_{m \times n} = U_{m \times m} \Sigma_{m \times n} V^*_{n \times n},$$

where $U$ is unitary with $UU^* = U^* U = I_{m \times m}$, $V$ is unitary with $VV^* = V^* V = I_{n \times n}$, and $\Sigma_{m \times n}$ is defined by

$$\Sigma_{ij} = \begin{cases} 
\sigma_i & i = j \text{ and } i \leq p \\
0 & \text{otherwise},
\end{cases}$$

where $p = \text{rank}(M_{m \times n})$, and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$. 

Singular Value Decomposition
The numbers $\sigma_1, \sigma_2, \ldots, \sigma_p$ are the singular values of $M$, while $U$ and $V$ result in from the spectral decomposition of $MM^*$ and $M^*M$, respectively. $MM^*$ is symmetric, non-negative of rank $p$, hence

$$MM^* = U\Lambda_{MM^*}U^*,$$

for some unitary $U$ and $\Lambda_{MM^*} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p, 0, \ldots, 0)_{m \times m}$ is a diagonal matrix with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$. $M^*M$ is also of rank $p$, symmetric with the same eigenvalues. As indeed if

$$MM^*x = \lambda x$$

then

$$M^*M(M^*x) = \lambda (M^*x).$$

Therefore we can write $M^*M = V\Lambda_{M^*M}V^*$ where $V$ is unitary and $\Lambda_{M^*M} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p, 0, \ldots, 0)_{n \times n}$, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$. 


From $M_{m \times m} = U \Sigma V^*$ we have

$$MM^* = U \Sigma V^* V \Sigma U^* = U \Sigma^2 U^*$$
$$M^* M = V \Sigma U^* U \Sigma V^* = V \Sigma^2 V^*,$$

hence, $\sigma_1^2 = \lambda_1, \sigma_2^2 = \lambda_2, \ldots, \sigma_p^2 = \lambda_p$.

The last $m - p$ eigenvectors in $U$ are an arbitrary set spanning the null-space of $MM^*$, and the last $n - p$ eigenvectors in $V$ are an arbitrary set spanning the null-space of $M^* M$.

$M$ can be written as $M =$

$$
\begin{pmatrix}
\bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_p
\end{pmatrix}_{m \times p}
\begin{pmatrix}
\sigma_1 \\
\sigma_2 \\
\vdots \\
\sigma_p
\end{pmatrix}_{m \times p}
\begin{pmatrix}
-\bar{v}_1^* \\
-\bar{v}_2^* \\
\vdots \\
-\bar{v}_p^*
\end{pmatrix}_{p \times n}
\sum_{k=1}^{p} \sigma_k \bar{u}_k \bar{v}_k^* = \sum_{k=1}^{p} \lambda_k^{1/2} \bar{u}_k \bar{v}_k^*$$
Returning to $\mathbb{H}$, we can write

$$
\mathbb{H}_{N \times N} = U_{N \times N} \Lambda_{N \times N}^{1/2} V_{N \times N}
$$

$$
= \sum_{k=1}^{p} \lambda_k^{1/2} \bar{u}_k \bar{v}_k^*. 
$$

In terms of spectral decomposition we have

$$
\bar{\psi}^{Data} = \mathbb{H} \bar{\psi}
$$

$$
= U \Lambda^{1/2} V^* \bar{\psi}
$$

$$
= \sum_{k=1}^{p} \lambda_k^{1/2} \bar{u}_k \bar{v}_k^* \bar{\psi}
$$

$$
= \sum_{k=1}^{p} \lambda_k^{1/2} \langle \bar{\psi}, \bar{v}_k \rangle \bar{u}_k.
$$
We would like to recover $\bar{\psi}$ from $\bar{\psi}^{Data}$. For $p < N$ the matrix $\mathbb{H}$ is not invertible. If $p = N$ then,

$$
\bar{\psi} = \mathbb{H}^{-1} \bar{\psi}^{Data}
= (U \Lambda^{1/2} V^*)^{-1} \bar{\psi}^{Data}
= V \Lambda^{-1/2} U^* \bar{\psi}^{Data}
= \sum_{k=1}^{N} \lambda_k^{-1/2} \bar{v}_k \bar{u}_k^* \bar{\psi}^{Data}
= \sum_{k=1}^{N} \lambda_k^{-1/2} \langle \bar{\psi}^{Data}, \bar{u}_k \rangle \bar{v}_k.
$$

When $p < N$ consider the modified $\mathbb{H}$

$$
\mathbb{H}(\epsilon) = U \left( \begin{array}{cccccc}
\lambda_1^{1/2} & & & & & \\
& \lambda_2^{1/2} & & & & \\
& & \ddots & & & \\
& & & \lambda_p^{1/2} & & \\
& & & & \epsilon & \\
& & & & & \epsilon \\
\end{array} \right) V^*.
$$
given $\mathbb{H}(\epsilon)$ we have

$$
\bar{\psi}(\epsilon) \equiv \mathbb{H}^{-1}(\epsilon)\bar{\psi}^{\text{Data}} = \sum_{k=1}^{p} \lambda^{-1/2} \langle \bar{\psi}^{\text{Data}}, \bar{u}_k \rangle \bar{v}_k + \sum_{k=p+1}^{N} \epsilon^{-1} \langle \bar{\psi}^{\text{Data}}, \bar{u}_k \rangle \bar{v}_k.
$$

Note however that

$$
\bar{\psi}^{\text{Data}} = \mathbb{H} \bar{\psi} = U \Lambda^{1/2} V^* \bar{\psi}
$$

$$
= U \begin{pmatrix}
& & & \\
& \lambda_1^{1/2} & & \\
& & \ddots & \\
& & & \lambda_p^{1/2} \\
& & & 0 \\
& & & \ddots \\
& & & & 0
\end{pmatrix} V^T \bar{\psi}
$$

$$
= \sum_{k=1}^{p} \lambda_k^{1/2} \langle \bar{\psi}, \bar{v}_k \rangle \bar{u}_k \in \text{span} (\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_p).
$$

Hence, for every $k \in \{p_1, \ldots, N\}$ we have $\langle \bar{\psi}^{\text{Data}}, \bar{u}_k \rangle = 0$. 
Therefore,

\[ \bar{\psi}(\epsilon) = H^{-1}(\epsilon)\bar{\psi}^{Data} \]

\[ = \sum_{k=1}^{p} \lambda_k^{-1/2} \langle \bar{\psi}^{Data}, \bar{u}_k \rangle \bar{v}_k + \sum_{k=p+1}^{N} \frac{1}{\epsilon} \cdot 0 \cdot \bar{v}_k \]

\[ = \sum_{k=1}^{p} \lambda_k^{-1/2} \langle \bar{\psi}^{Data}, \bar{u}_k \rangle \bar{v}_k, \]

independent of \( \epsilon \)!

That is, for any \( \epsilon \) we choose, \( \bar{\psi}(\epsilon) \equiv \hat{\psi} = H^{-1}(\epsilon)\bar{\psi}^{Data} \) is the same.
Thus, we simply select

\[
\mathbb{H}^{-1} \equiv V \begin{pmatrix}
\lambda_1^{-1/2} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{pmatrix} U^*,
\]

to be the inverse operator applied to \( \bar{\psi}^{Data} \) to recover \( \hat{\psi} \) as an estimate of \( \bar{\psi} \). This selection is as if we take \( \epsilon \to 0 \).
We obtained

\[ \hat{\psi} = \mathbb{H}^{-1} \bar{\psi}_{Data} \]

\[ = V \begin{pmatrix} \lambda_1^{-1/2} & \cdots & \cdots & \lambda_p^{-1/2} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} U^* \bar{v}_{Data} \]

\[ = \sum_{k=1}^{p} \lambda_k^{-1/2} \langle \bar{\psi}_{Data}, \bar{u}_k \rangle \bar{v}_k \in \text{span}(\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_p). \]
What is the MSE if recovering $\hat{\psi}$ rather than $\bar{\psi}$?

Let us represent $\bar{\psi}$ in the basis $V = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_N)$, that is,

$$\bar{\psi} = VV^* \bar{\psi} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_N) \begin{pmatrix} \langle \bar{v}_1, \bar{\psi} \rangle \\ \vdots \\ \langle \bar{v}_n, \bar{\psi} \rangle \end{pmatrix},$$
Then, $\mathbb{H} \tilde{\psi}$ can be written as

$$
\psi^{Data} = \mathbb{H} \tilde{\psi} = \mathbb{H} V V^* \tilde{\psi} =
U \begin{pmatrix}
\lambda_1 \\
\vdots \\
\lambda_p
\end{pmatrix}
\begin{pmatrix}
| & | & | & | & | \\
| & | & | & | & | \\
V^* & & & & \\
| & | & | & | & | \\
| & | & | & | & | \\
\langle \bar{v}_1, \tilde{\psi} \rangle & \cdots & \langle \bar{v}_n, \tilde{\psi} \rangle
\end{pmatrix}
V V^* \tilde{\psi} =
U \begin{pmatrix}
\lambda_1 \\
\vdots \\
\lambda_p
\end{pmatrix}
\begin{pmatrix}
| & | & | & | & | \\
| & | & | & | & | \\
\lambda_1 & & & & \\
| & | & | & | & | \\
| & | & | & | & | \\
\langle \bar{v}_1, \tilde{\psi} \rangle & \cdots & \langle \bar{v}_n, \tilde{\psi} \rangle
\end{pmatrix} =
\sum_{k=1}^{p} \lambda_k \langle \bar{\psi}, \bar{v}_k \rangle \bar{u}_k.
$$
Therefore, if we write

\[
\bar{\psi} = \sum_{k=1}^{p} \langle \bar{\psi}, \bar{v}_k \rangle \bar{v}_k + \sum_{k=p+1}^{N} \langle \bar{\psi}, \bar{v}_k \rangle \bar{v}_k,
\]

we see that \( \hat{\psi} \) is simply the part of \( \bar{\psi} \) which is in the subspace spanned by \( \{ (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_p) \} \). Writing

\[
\hat{\psi} = \sum_{k=1}^{p} \lambda_k^{-1/2} \langle \bar{\psi}^{Data}, \bar{u}_k \rangle \bar{u}_v,
\]

and expanding, we get

\[
\hat{\psi} = \sum_{k=1}^{p} \lambda_k^{-1/2} \left( \sum_{l=1}^{p} \lambda_l^{1/2} \langle \bar{\psi}, \bar{v}_l \rangle \bar{u}_l, \bar{u}_k \right) \bar{v}_k
\]

\[
= \sum_{k=1}^{p} \lambda_k^{-1/2} \left( \sum_{l=1}^{p} \lambda_l^{1/2} \langle \bar{\psi}, \bar{v}_l \rangle \langle \bar{u}_l, \bar{u}_k \rangle \delta_{kl} \right) \bar{v}_k
\]

\[
= \sum_{k=1}^{p} \lambda_k^{-1/2} \lambda_k^{1/2} \langle \bar{\psi}, \bar{v}_k \rangle \cdot 1 \cdot \bar{v}_k = \sum_{k=1}^{p} \langle \bar{\psi}, \bar{v}_k \rangle \bar{v}_k.
\]
This can be easily seen from

\[
\hat{\psi} = H^{-1} \tilde{\psi}_{Data}
\]

\[
= V \left( \begin{array}{cccc}
\lambda_1^{-1/2} & \cdots & \lambda_p^{-1/2} & 0 \\
\lambda_1^{-1/2} & \cdots & \lambda_p^{-1/2} & 0 \\
1 & \cdots & 1 & 0 \\
0 & \cdots & 0 & 0
\end{array} \right) U^* U \left( \begin{array}{cccc}
\lambda_1^{1/2} & \cdots & \lambda_p^{1/2} & 0 \\
\lambda_1^{1/2} & \cdots & \lambda_p^{1/2} & 0 \\
1 & \cdots & 1 & 0 \\
0 & \cdots & 0 & 0
\end{array} \right) V^* \tilde{\psi}.
\]

The pseudo inverse recovers the part of \( \tilde{\psi} \) which is in the span of \( \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_p \). The part that belongs to the span of \( \tilde{v}_{p+1}, \ldots, \tilde{v}_N \) can not be recovered since any vector in that subspace is mapped to zero by \( H \). The span of \( \{ \tilde{v}_{p+1}, \ldots, \tilde{v}_N \} \) is the null space of \( H \), or the kernel of \( H \) defined as \( \text{NullSpace} H = \{ \tilde{x} | H\tilde{x} = 0 \} \).

Hence, given \( \tilde{\psi} \) the pseudo-inverse of \( H \) recovers \( \hat{\psi} \), which is the part of \( \tilde{\psi} \) which is not in the null space of \( H \).
Reconstruction error

\[ \| \bar{\psi} - \hat{\psi} \|^2 = \| \sum_{k=1}^{p} \langle \bar{\psi}, \bar{v}_k \rangle \bar{v}_k \|^2 \]

\[ = \sum_{k=1}^{p} \langle \bar{\psi}, \bar{v}_k \rangle^2, \]

where \( \hat{\psi} = \sum_{k=1}^{p} \langle \bar{\psi}, \bar{v}_k \rangle \bar{v}_k. \)

Again, \( \bar{\psi}^{Data} = \mathbb{H} \hat{\psi} = \mathbb{H} \bar{\psi} \), and we could add to \( \hat{\psi} \) any vector in the null-space of \( \mathbb{H} \) and obtain the same \( \bar{\psi}^{Data} \). Indeed,

\[ \mathbb{H} \left( \hat{\psi} + \sum_{k=p+1}^{N} \alpha_k \bar{v}_k \right) = \mathbb{H} (\hat{\psi}) = \mathbb{H} (\bar{\psi}). \]
Note that

\[
\left\| \hat{\psi} + \sum_{k=p+1}^{N} \alpha_k \bar{v}_k \right\|^2 = \|\hat{\psi}\|^2 + \sum_{k=p+1}^{N} \alpha_k^2.
\]

Thus, \(\hat{\psi}\) is the shortest vector obeying \(H\hat{\psi} = \bar{\psi}^{Data}\). The pseudo-inverse solves the optimization problem

\[
\text{minimize} \quad \|\bar{\psi}\|^2 \equiv \bar{\psi}^* \bar{\psi}
\]

subject to \(H\bar{\psi} = \bar{\psi}^{Data}\).

In the sequel we will see how many other signal restoration problems are posed as optimization problems.
Discrete Inverse filtering (the circulant case)

Given $H$ a circulant matrix. That is $\mathcal{H}(\cdot)$ a linear shift invariant (LSI) operator. Then,

$$H_{LSI} = (DFT)^* \begin{pmatrix} \lambda_0^H & 0 & & \\ & \lambda_1^H & & \\ & & \ddots & \\ & & & \lambda_{N-1}^H \end{pmatrix} (DFT),$$

where

$$\begin{pmatrix} \lambda_0^H \\ \lambda_1^H \\ \vdots \\ \lambda_{N-1}^H \end{pmatrix} = (DFT) \begin{pmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{pmatrix}.$$
Note that this is not the same as the SVD of $H_{LSI}$ which is based on

$$HH^* = (DFT)^* \Lambda^H (DFT)(DFT)^* \Lambda^H (DFT) = (DFT)^* \Lambda^H \Lambda^H (DFT)$$

and

$$H^*H = (DFT)^* \Lambda^{H*} (DFT)(DFT)^* \Lambda^H (DFT) = (DFT)^* \Lambda^{H*} \Lambda^H (DFT).$$

and

$$\Lambda^{H*} \Lambda^H = \Lambda^H \Lambda^{H*} = \begin{pmatrix} |\lambda_0^H|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_{N-1}^H|^2 \end{pmatrix}$$

contains only positive values.
Here we can write

\[ \mathcal{H}_{LSI} = (DFT)^*(PH) \begin{pmatrix} |\lambda_0^H| & 0 \\ 0 & |\lambda_1^H| \\ & \ddots \\ & 0 & |\lambda_{N-1}^H| \end{pmatrix} (DFT), \]

where \((PH) \equiv \begin{pmatrix} e^{i\alpha_0} & 0 \\ 0 & e^{i\alpha_1} \\ & \ddots \\ & 0 & e^{i\alpha_{N-1}} \end{pmatrix}, \)

where \(\lambda_k^H \equiv |\lambda_k| e^{i\alpha_k},\)

or

\[ \mathcal{H}_{LSI} = (DFT)^* \begin{pmatrix} |\lambda_0^H| & 0 \\ 0 & |\lambda_1^H| \\ & \ddots \\ & 0 & |\lambda_{N-1}^H| \end{pmatrix} (PH)(DFT), \]

are the possible SVD’s of \(\mathcal{H}_{LSI}.\)
Clearly,

\[
((DFT)^*(PH))((DFT)^*(PH)) = (DFT)^*(PH)(PH)^*(DFT) = I,
\]

and

\[
\]

Note that circulants commute, hence \( HHH^* = H^*H \). The SVD is non-unique, as the matrices \( U \) and \( V \) in \( M_{N \times N} = U\Lambda^{1/2}V^* \) can always be modified to \( U PH_{\alpha_1,\alpha_2,\ldots,\alpha_N} \), and \( V PH_{\alpha_1,\alpha_2,\ldots,\alpha_N} \) to get the same result \( M_{N \times N} = U \underbrace{PH\Lambda^{1/2}PH^*}_{\Lambda^{1/2}} V^* \).
Back to our basic problem where

$$\bar{\psi}^{Data} = \bar{\mathbb{H}} \bar{\psi} = (DFT)^* \Lambda^H (DFT) \bar{\psi},$$

which yields,

$$(DFT) \bar{\psi}^{Data} = \Lambda^H (DFT) \bar{\psi}.$$

In the transform domain, to recover $\bar{\psi}$ we need to invert the diagonal matrix $\text{diag}(\lambda_0^H, \lambda_2^H, \ldots, \lambda_{N-1}^H)$ and multiply by the $(DFT)$ applied to the vector $(h_0 \ h_1 \cdots \ h_{N-1})^*$, the discrete impulse response of $\mathcal{H}_{LSI}(\cdot)$. 
When some $\lambda_k^H = 0$ we will have a problem, and the (deconvolution) inversion process would be impossible to perform. Yet, places where $\lambda_k^H = 0$ we also have that $(DFT)\tilde{\psi}^{Data}$ is zero. Denote,

$$\begin{pmatrix} \tilde{\psi}_0^{Data} \\ \tilde{\psi}_1^{Data} \\ \vdots \\ \tilde{\psi}_{N-1}^{Data} \end{pmatrix} = (DFT)\tilde{\psi}^{Data},$$

and by

$$\begin{pmatrix} \tilde{\psi}_0 \\ \tilde{\psi}_1 \\ \vdots \\ \tilde{\psi}_{N-1} \end{pmatrix} = (DFT)\tilde{\psi},$$

we have in the transform domain $\tilde{\psi}_k^{Data} = \lambda_k^H \tilde{\psi}_k$. 
If $\lambda_k^H = 0$ we know that $\tilde{\psi}_k^{Data} = 0$ as well, else, we have for such $k$ that $\lambda_k^H \neq 0$.

$$\tilde{\psi}_k = \frac{1}{\lambda_k^H} \tilde{\psi}_k^{Data} = \frac{1}{|\lambda_k^H|} e^{-i\theta(\lambda_k^H)} \tilde{\psi}_k^{Data},$$

where we used the definition of a phase by which $\lambda_k^H = |\lambda_k^H| e^{i\theta(\lambda_k^H)}$. The result is

$$\tilde{\psi}_k = \begin{cases} 
\frac{1}{|\lambda_k^H|} e^{-i\theta(\lambda_k^H)} \tilde{\psi}_k^{Data} & \text{if } |\lambda_k^H| \neq 0 \\
0 & \text{if } |\lambda_k^H| = 0.
\end{cases}$$

This is the inverse filter solution for deconvolution in the Fourier domain.
Example

Sliding window (smoothing) operator, where $H$ is a circulant matrix with first row $(1 \ldots 1 0 \cdots 0)$ with $m$ ones and $N - m$ zeros.

Hence, $\overline{\psi}_{Data} = (1 0 0 \cdots 1 \cdots 1)^{m} \ast (\psi_0 \psi_1 \cdots \psi_{N-1}).$

We have the following filter

$$
\begin{pmatrix}
\lambda_0^H \\
\lambda_1^H \\
\vdots \\
\lambda_{N-1}^H
\end{pmatrix}
= (DFT)^{*}
\begin{pmatrix}
1 \\
\vdots \\
1 \\
0
\end{pmatrix}
= \cdots
$$

workout at home.

Show that if $N = 2^n$ and $m = 2$ then $e^{i\frac{2\pi}{2^n}k^2} = 1$ if $2k = 2^n \Rightarrow k = 2^{n-1}$.

Therefore, $\lambda_{2^{n-1}} = \lambda_{N/2} = 0$ and the matrix $H_{LSI}$ is not invertible. In general, if $k = (\text{integer}/m)N$ we will have $|\lambda_k| = 0$. In these cases we will have to apply the pseudo-inverse filter.