Signal, Image and Data Processing  (236200)

Tutorial 5

Signal Representation using Families of Orthonormal Functions
Reminder:
A Family of Orthonormal Functions

A set of orthonormal (continuous and defined over [0,1]) functions

\{\psi_1(t), \psi_2(t), \ldots, \psi_N(t), \ldots\}

satisfies the orthonormality condition

\[ \langle \psi_i(t), \psi_j(t) \rangle = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } i \neq j \end{cases} \]

where the inner-product is defined as

\[ \langle \psi_i, \psi_j \rangle = \int_0^1 \psi_i(t) \psi_j(t) dt \]
Reminder:
Representing Signals using Orthonormal Functions

We represent a given signal $\varphi(t)$ by approximating it using a set of $N$ orthonormal functions:

$$\hat{\varphi}(t) = \sum_{i=1}^{N} \varphi_i \cdot \psi_i(t)$$

where $\{\varphi_i\}_{i=1}^{N}$ are the representation coefficients.

What are the optimal representation-coefficients? (for the MSE criterion)
Reminder:
Representing Signals using Orthonormal Functions

What are the optimal representation-coefficients?

The MSE for $N$-term approximation is

$$MSE(N) = \int_0^1 (\varphi(t) - \hat{\varphi}(t))^2 dt = \int_0^1 \left( \varphi(t) - \sum_{i=1}^N \varphi_i \cdot \psi_i(t) \right)^2 dt$$

The optimal $\{\varphi_i\}_{i=1}^N$ are given by demanding

$$\frac{\partial MSE(N)}{\partial \varphi_i} = 0 \quad \text{for} \quad i = 1, \ldots, N$$

that results in $\varphi_i^{opt} = \langle \varphi(t), \psi_i(t) \rangle = \int_0^1 \varphi(t) \psi_i(t) dt$
Reminder:
Representing Signals using Orthonormal Functions

What are the optimal representation-coefficients?

Proof:

The optimal \( \{ \varphi_i \}_{i=1}^{N} \) are given by demanding \( \frac{\partial \text{MSE}(N)}{\partial \varphi_i} = 0 \) for \( i = 1, ..., N \)

\[
\frac{\partial}{\partial \varphi_i} \int_0^1 \left( \varphi(t) - \sum_{j=1}^{N} \varphi \cdot \psi_j(t) \right)^2 dt = \int_0^1 \frac{\partial}{\partial \varphi_i} \left\{ \left( \varphi(t) - \sum_{j=1}^{N} \varphi \cdot \psi_j(t) \right)^2 \right\} dt = \\
= -2 \int_0^1 \psi_i(t) \cdot \left( \varphi(t) - \sum_{j=1}^{N} \varphi \cdot \psi_j(t) \right) dt = -2 \int_0^1 \psi_i(t) \varphi(t) dt + 2 \sum_{j=1}^{N} \int_0^1 \psi_i(t) \psi_j(t) dt \\
= -2 \int_0^1 \psi_i(t) \varphi(t) dt + 2 \varphi_i
\]

Demanding equality to zero:

\[
-2 \int_0^1 \psi_i(t) \varphi(t) dt + 2 \varphi_i = 0 \quad \rightarrow \quad \varphi_{i}^{opt} = \int_0^1 \psi_i(t) \varphi(t) dt
\]
Reminder:
Representing Signals using Orthonormal Functions

We can further develop the MSE for $N$-term approximation:

$$MSE(N) = \int_0^1 \left( \phi(t) - \sum_{i=1}^N \phi_{i, opt} \cdot \psi_i(t) \right)^2 dt$$

$$= \int_0^1 \left[ \phi^2(t) - 2\phi(t) \cdot \sum_{i=1}^N \phi_{i, opt} \cdot \psi_i(t) + \left( \sum_{i=1}^N \phi_{i, opt} \cdot \psi_i(t) \right)^2 \right] dt$$

$$= \int_0^1 \phi^2(t) dt - \int_0^1 2\phi(t) \cdot \sum_{i=1}^N \phi_{i, opt} \cdot \psi_i(t) dt + \int_0^1 \left( \sum_{i=1}^N \phi_{i, opt} \cdot \psi_i(t) \right)^2 dt$$

$$= \int_0^1 \phi^2(t) dt - 2 \sum_{i=1}^N (\phi_{i, opt})^2 + \sum_{i=1}^N (\phi_{i, opt})^2$$

$$= \int_0^1 \phi^2(t) dt - \sum_{i=1}^N (\phi_{i, opt})^2$$
The Perspective of Generalized Sampling
Standard Sampling

Let us consider the standard sampling functions:

\[
\psi^s_i(t) = \begin{cases} 
\sqrt{N} & \text{for } t \in \left[\frac{i-1}{N}, \frac{i}{N}\right) \\
0 & \text{otherwise}
\end{cases}
\]

Properties of these functions:

- **Tile the entire domain** [0,1).
- Have local non-overlapping supports.
- **Constant valued** over the supports.
The Perspective of Generalized Sampling

Standard Sampling

Show orthonormality:

For $i \neq j$:
\[
\langle \psi_i^s(t), \psi_j^s(t) \rangle = \int_0^1 \psi_i^s(t)\psi_j^s(t)dt
\]
\[
= \int_{i-1}^{\frac{N}{i}} \sqrt{N} \cdot 0 \ dt + \int_{\frac{N}{i-j-1}}^N 0 \cdot \sqrt{N} dt = 0
\]

For $i = j$:
\[
\langle \psi_i^s(t), \psi_i^s(t) \rangle = \int_0^1 \psi_i^s(t)\psi_i^s(t)dt = \int_{\frac{N}{i-1}}^{\frac{N}{i}} \sqrt{N} \cdot \sqrt{N} dt = 1
\]
The Perspective of Generalized Sampling
Standard Sampling

The general formulas reduce to the previous results of the standard case:

- For example, the **optimal coefficients** are
  the **normalized averages** over the sampling intervals:

\[
\phi_{i}^{opt} = \langle \varphi(t), \psi_{i}(t) \rangle = \int_{0}^{1} \varphi(t)\psi_{i}(t)dt = \int_{\frac{i}{N}}^{\frac{i+1}{N}} \sqrt{N} \cdot \varphi(t)dt
\]

\[
= \frac{1}{\sqrt{N}} \left( \frac{1}{1/N} \int_{\frac{i-1}{N}}^{\frac{i}{N}} \varphi(t)dt \right)
\]
Standard Sampling: The Implications of Locality

The standard-family functions have **local non-overlapping** supports, therefore, **a subset of the coefficients** allows **reconstruction only over parts of the signal’s domain**:

The reconstructed signal is not defined over the intervals corresponding to the omitted coefficients.
Unitary Matrices

The orthonormality condition for a square matrix \( U \) of size \( N \times N \):

- For general complex-valued matrices: \( U^*U = I \)
  where \( * \) is the conjugate-transpose operator.
- For real-valued matrices: \( U^T U = I \)
  where the conjugate-transpose reduces to transpose only.

The above implies that the matrix columns are orthonormal:

\[
U_i^*U_j = \begin{cases} 
1 & \text{for } i = j \\
0 & \text{for } i \neq j 
\end{cases}
\]

where \( U_i \) is the \( i^{th} \) column of the matrix \( U \).

(similarly, the rows of \( U \) are also orthonormal).

I.e., \( \{U_i\}_{i=1}^N \) are \( N \) linearly-independant vectors of length \( N \).
Therefore, they form a basis of \( \mathbb{R}^N \) (in the real case) or \( \mathbb{C}^N \) (in the complex case).
Constructing Orthonormal Functions from a Unitary Matrix

Let us consider a real-valued unitary matrix $U$ of size $N \times N$, formed from $N$ column vectors $\{u_i\}_{i=1}^N$.

The direct transform of some signal vector $x$ is considered here as $U^T x$ (note the transpose).

An orthonormal set of $N$ continuous functions can be constructed as:

$$\psi_i(t) = \sum_{k=1}^{N} u_i^{(k)} \cdot \psi^s_k(t) \quad i = 1, \ldots, N$$

where, $u_i^{(k)}$ is the $k^{th}$ element of the $i^{th}$ column-vector $u_i$.

$\psi^s_k(t)$ is the $k^{th}$ function of the standard family (for $N$ samples).

The orthonormality of the columns, $\{u_i\}_{i=1}^N$, leads to orthonormality of the functions $\{\psi_i(t)\}_{i=1}^N$.

(as proved in the lecture).
The Hadamard Basis

The Hadamard matrices are recursively defined as follows:

\[ H_1 = 1 \]
\[ H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \]
\[ H_{2^n} = H_2 \otimes H_{2^{n-1}} = \frac{1}{\sqrt{2}} \begin{bmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{bmatrix} \quad \text{for } n = 2, 3, \ldots \]

Note that \( H_{2^n} \) is a \( 2^n \times 2^n \) matrix.

Examples:

\[ H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \]
\[ H_8 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix} \]
The Walsh-Hadamard Basis

**Walsh-Hadamard** matrices are the Hadamard matrices with rows ordered according to their **sequency** (the number of sign changes).

Example:

\[
H_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}
\]

The sequency is: 0, 3, 1, 2

Transition to Walsh matrix via row reordering:

\[
W_4 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}
\]

The sequency is: 0, 1, 2, 3

The 4 **orthonormal continuous functions** corresponding to the above Walsh-Hadamard matrix \(W_4\) are:

\[
\psi_i(t) = \sum_{k=1}^{4} w_{k,i} \cdot \psi_k^s(t) \quad i = 1, \ldots, 4
\]

where \(\psi_k^s(t)\) are the standard sampling functions, and \(w_{k,i}\) is the \((k, i)\) element of the matrix \(W_4\).
Exercise: Comparing Sampling using the Standard and Hadamard Functions

The signal \( \varphi(t) = \begin{cases} t, & \text{for } t \in [0,1) \\ 0, & \text{otherwise} \end{cases} \)

is sampled using the standard sampling function to have \( N=4 \) samples.

Calculate the representation coefficients and demonstrate the \( K \)-term approximation and its error.
Exercise: Comparing Sampling using the Standard and Hadamard Functions

Consider the standard sampling functions \((N=4)\).
What are the optimal coefficients?

Let’s define the sampling-interval size \(\Delta = \frac{1}{N} = \frac{1}{4}\).

The coefficients are the normalized interval averages which in this case (where \(\varphi(t)\) is linear) are the normalized interval centers:

\[
\varphi_i^s = \frac{1}{\sqrt{N}} \left( \frac{1}{1/N} \int_{i-1/N}^{i/N} \varphi(t) \, dt \right) = \frac{1}{\sqrt{N}} \Delta \cdot \left( i - \frac{1}{2} \right)
\]

\[
\begin{align*}
\varphi_1^s &= \frac{1}{16} \\
\varphi_2^s &= \frac{3}{16} \\
\varphi_3^s &= \frac{5}{16} \\
\varphi_4^s &= \frac{7}{16}
\end{align*}
\]

Recall the normalization included in the coefficient values.
Exercise: Comparing Sampling using the Standard and Hadamard Functions

The K-term approximation \((K = 1, \ldots, N)\) for the standard sampling functions \((N=4)\):

1 term
\[
\varphi_s^1 = \frac{7}{16}
\]

2 terms
\[
\varphi_s^2 = \frac{5}{16}
\]

3 terms
\[
\varphi_s^3 = \frac{3}{16}
\]

4 terms
\[
\varphi_s^4 = \frac{1}{16}
\]

\[
MSE(1) = \frac{1}{3} - \left(\frac{7}{16}\right)^2 = 0.1419
\]

\[
MSE(2) = 0.0443
\]

\[
MSE(3) = 0.0091
\]

\[
MSE(4) = 0.0052
\]
Exercise: Comparing Sampling using the Standard and Hadamard Functions

Now we represent the signal using the \textit{Walsh-Hadamard} function ($N = 4$). What are the coefficients and the MSE?

For $i = 1, \ldots, 4$: \[
\varphi_i^w = \langle \varphi(t), \psi_i^w(t) \rangle = \int_0^1 t \psi_i^w(t) dt
\]

We will use the following auxiliary-integral: \[
I(a, b) = \int_a^b t dt = \frac{t^2}{2} \bigg|_a^b = \frac{b^2 - a^2}{2}
\]

\[
\varphi_1^w = \langle \varphi(t), \psi_1^w(t) \rangle = \int_0^1 t dt = I(0, 1) = \frac{1}{2}
\]

\[
\varphi_2^w = \langle \varphi(t), \psi_2^w(t) \rangle = I\left(0, \frac{1}{2}\right) - I\left(\frac{1}{2}, 1\right) = \frac{1}{8} - \frac{3}{8} = -\frac{1}{4}
\]
Exercise: Comparing Sampling using the Standard and Hadamard Functions

\[ \varphi_3^w = \langle \varphi(t), \psi_3^w(t) \rangle = I\left(0, \frac{1}{4}\right) - I\left(\frac{1}{4}, \frac{3}{4}\right) + I\left(\frac{3}{4}, 1\right) \]

\[ = \frac{1}{32} - \frac{8}{32} + \frac{7}{32} = 0 \]

\[ \varphi_4^w = \langle \varphi(t), \psi_4^w(t) \rangle = I\left(0, \frac{1}{4}\right) - I\left(\frac{1}{4}, \frac{1}{2}\right) + I\left(\frac{1}{2}, \frac{3}{4}\right) - I\left(\frac{3}{4}, 1\right) \]

\[ = \frac{1}{32} - \frac{3}{32} + \frac{5}{32} - \frac{7}{32} = -\frac{1}{8} \]
Exercise: Comparing Sampling using the Standard and Hadamard Functions

The K-term approximation for the *Walsh-Hadamard* sampling functions ($N=4$):

1 term

\[ \varphi_1^w = \frac{1}{2} \]

2 terms

\[ \varphi_2^w = -\frac{1}{4} \]

3 terms

\[ \varphi_3^w = 0 \]

4 terms

\[ \varphi_4^w = -\frac{1}{8} \]

The 4-term approximation equals to the 3-term one.

\[ MSE(4) = MSE(3) \]

\[ MSE(1) = \frac{1}{3} - \left(\frac{1}{2}\right)^2 = 0.0833 \]

\[ MSE(2) = 0.0208 \]

\[ MSE(3) = 0.0052 \]

\[ MSE(K) = \int_0^1 \varphi^2(t) dt - \sum_{i=1}^{K} (\varphi_i^w)^2 \]
Exercise: Comparing Sampling using the Standard and Hadamard Functions

Comparing the MSE curves shows the superiority of approximation using the Walsh-Hadamard functions:

- Note that in both cases the error for $K = N$ is the same (why?).
Exercise #2: Representation with Degraded Coefficients

Consider $N$ orthonormal functions $\{\beta_i(t)\}_{i=1}^N$, where $\beta_i: [0,1) \rightarrow \mathbb{R}$.

The input signal is $\varphi: [0,1) \rightarrow \mathbb{R}$.

A signal $\varphi(t)$ is optimally represented based on $\{\beta_i(t)\}_{i=1}^N$ and the corresponding $N$ optimal-coefficients $\{\varphi_i^{opt}\}_{i=1}^N$.

The optimal coefficients, $\{\varphi_i^{opt}\}_{i=1}^N$, are degraded by some scalar degradation procedure in the form of the function $h: \mathbb{R} \rightarrow \mathbb{R}$, resulting in the degraded coefficients:

$$\varphi_i^D = h(\varphi_i^{opt}), \quad i = 1, ..., N$$

The squared error due to the coefficient degradation (with respect to the optimal coefficients) is

$$\mathcal{E}_D^2 = \sum_{i=1}^{N} (\varphi_i^{opt} - \varphi_i^D)^2$$

236200, CS Department, Technion
Exercise #2: Representation with Degraded Coefficients

The signal reconstructed from the degraded coefficients is

$$\hat{\phi}^D(t) = \sum_{i=1}^{N} \phi^D_i \beta_i(t)$$

Develop (in detail) the MSE expression for $\hat{\phi}^D(t)$ with respect to the input signal $\phi(t)$, i.e.,

$$MSE_{\beta D} = \int_{0}^{1} (\phi(t) - \hat{\phi}^D(t))^2 dt$$

and show that

$$MSE_{\beta D} = MSE_{\beta} + \mathcal{E}_D^2$$

where $MSE_{\beta}$ is the approximation MSE using the optimal coefficients, i.e.,

$$MSE_{\beta} = \int_{0}^{1} \phi^2(t) dt - \sum_{i=1}^{N} (\phi^{opt}_i)^2$$
Exercise #2: 
Representation with Degraded Coefficients

Solution:

The signal approximation with the degraded coefficients is

\[ \hat{\varphi}^D(t) = \sum_{i=1}^{N} \varphi_i^D \beta_i(t) \]

and the corresponding MSE is

\[ MSE_{\beta_D} = \int_0^1 \left( \varphi(t) - \sum_{i=1}^{N} \varphi_i^D \beta_i(t) \right)^2 dt \]

Let us develop the \( MSE_{\beta_D} \) expression to show the requested decomposition (see next slide).
Exercise #2: Representation with Degraded Coefficients

\[ MSE_{\beta_D} = \int_0^1 \left( \varphi(t) - \sum_{i=1}^{N} \varphi_i^D \beta_i(t) \right)^2 dt = \int_0^1 \left( \varphi^2(t) - 2\varphi(t) \sum_{i=1}^{N} \varphi_i^D \beta_i(t) + \left( \sum_{i=1}^{N} \varphi_i^D \beta_i(t) \right)^2 \right) dt \]

\[ = \int_0^1 \varphi^2(t) dt - 2 \sum_{i=1}^{N} \varphi_i^D \left( \int_0^1 \varphi(t) \beta_i(t) dt \right) + \sum_{i=1}^{N} \sum_{l=1}^{N} \varphi_i^D \varphi_l^D \left( \int_0^1 \beta_i(t) \beta_l(t) dt \right) \]

\[ = \int_0^1 \varphi^2(t) dt - 2 \sum_{i=1}^{N} \varphi_i^D \varphi_i^{opt} + \sum_{i=1}^{N} (\varphi_i^D)^2 \]

\[ = \int_0^1 \varphi^2(t) dt - \sum_{i=1}^{N} (\varphi_i^{opt})^2 + \sum_{i=1}^{N} (\varphi_i^{opt})^2 - 2 \sum_{i=1}^{N} \varphi_i^D \varphi_i^{opt} + \sum_{i=1}^{N} (\varphi_i^D)^2 \]

\[ = \int_0^1 \varphi^2(t) dt - \sum_{i=1}^{N} (\varphi_i^{opt})^2 + \sum_{i=1}^{N} (\varphi_i^{opt} - \varphi_i^D)^2 \]

(Completing the square)
Exercise #2: Representation with Degraded Coefficients

Recall that the signal-representation MSE with non-degraded coefficients is

\[ MSE_\beta = \int_0^1 \varphi^2(t) dt - \sum_{i=1}^N (\varphi_{i opt}^2) \]

and that the (total) squared error of coefficient degradation is

\[ \mathcal{E}_D^2 = \sum_{i=1}^N (\varphi_{i opt} - \varphi_i^D)^2 \]

Therefore, we got that the MSE of representing the signal \( \varphi(t) \) with the degraded coefficients is

\[ MSE_{\beta D} = MSE_\beta + \mathcal{E}_D^2 \]

Showing the separability of the generalized-sampling MSE and the degradation squared error.