The Haar Transform Matrix

The N Haar functions can be sampled at $t = m/N$, where $m = 0, \cdots, N-1$ to form an N by N matrix for discrete Haar transform. For example, when $N = 2$, we have

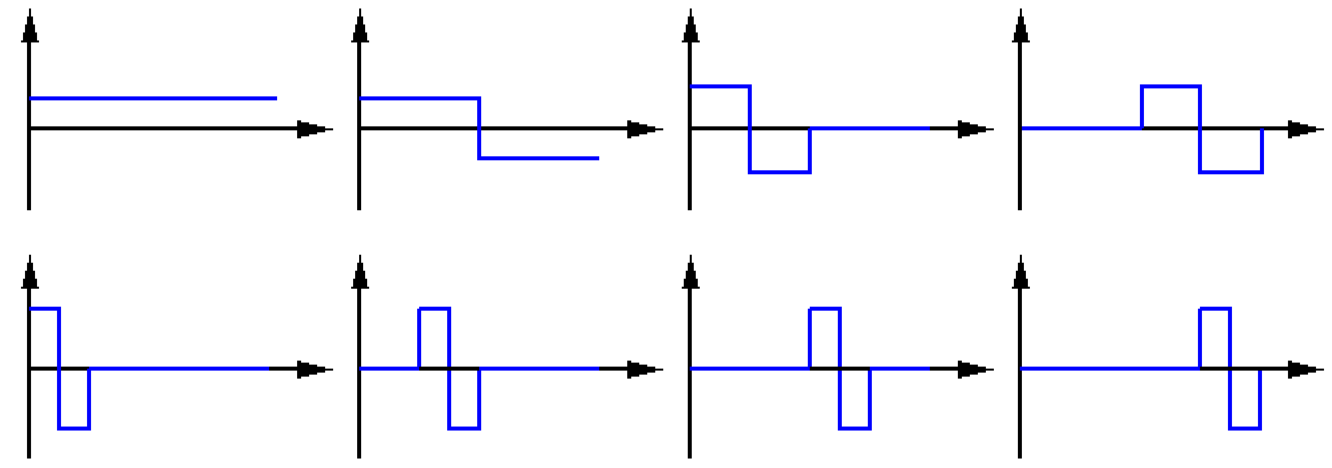
$$\mathbf{H}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

when $N = 4$, we have

$$\mathbf{H}_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

and when $N = 8$

$$\mathbf{H}_8 = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix} \begin{matrix} \varphi_0(t) \\ \psi_0(t) \\ \psi_{1,0}(t) \\ \psi_{1,1}(t) \\ \psi_{2,0}(t) \\ \psi_{2,1}(t) \\ \psi_{2,2}(t) \\ \psi_{2,3}(t) \end{matrix}$$



We see that all Haar functions $h_k(t),\ (k > 0)$ contains a single prototype shape composed of a square wave and its negative version, and the parameters

- p specifies the magnitude and width (or scale) of the shape;
- q specifies the position (or shift) of the shape.

Note that the functions $h_k(t)$ of Haar trnasform can represent not only the details in the signal of different scales (corresponding to different frequencies) but also their locations in time.

The Haar transform matrix is real and orthogonal:

$$\mathbf{H} = \mathbf{H}^*, \qquad \mathbf{H}^{-1} = \mathbf{H}^T, \quad \text{i.e.} \quad \mathbf{H}^T\mathbf{H} = \mathbf{I}$$

where $\mathbf{I}$ is identity matrix. For example, when $N = 4$,

$$\mathbf{H}_4^{-1}\mathbf{H}_4 = \mathbf{H}_4^T\mathbf{H}_4 = \frac{1}{4}\begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In general, an N by N Haar matrix can be expressed in terms of its row vectors:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_0^T \\ \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_{N-1}^T \end{bmatrix}, \qquad \mathbf{H}^{-1} = \mathbf{H}^T = [\mathbf{h_0}, \cdots, \mathbf{h_{N-1}}]$$

where $\mathbf{h}_n^T$ is the nth row vector of the matrix. The Haar tansform of a given signal vector $\mathbf{x} = [x[0], \cdots, x[N-1]]^T$ is

$$\mathbf{X} = \mathbf{H}\mathbf{x} = [\mathbf{h_0}, \cdots, \mathbf{h_{N-1}}]\mathbf{x}$$

with the n-th component of $\mathbf{X}$ being

$$X[n] = \mathbf{h}_n^T \mathbf{x} \qquad (n = 0, \cdots, N-1)$$

which is the nth transform coefficient, the projection of the signal vector $\mathbf{x}$ onto the n-th row vector of the transform $\mathbf{H}$ matrix. The inverse transform is

$$\mathbf{x} = \mathbf{H}^{-1}\mathbf{X} = \mathbf{H}^T\mathbf{X} = [\mathbf{h_0}, \cdots, \mathbf{h}_{N-1}] \begin{bmatrix} X[0] \\ \vdots \\ X[N-1] \end{bmatrix} = \sum_{n=0}^{N-1} X[n]\mathbf{h}_n$$

i.e., the signal is expressed as a linear combination of the row vectors of $\mathbf{H}$.

Comparing this Haar transform matrix with all transform matrices previously discussed (e.g., Fourier transform, cosine transform, Walsh-Hadamard transform), we see an essential difference. The row vectors of all previous trnasform methods represent different frequency (or sequency) components, including zero frequency or the average or DC component (first row $n = 0$), and the progressively higher frequencies (sequencies) in the subsequent rows ($n = 1, \cdots, N-1$). However, the row vectors in Haar transform matrix represent progressively smaller scales (narrower width of the square waves) and their different positions. It is the capability to represent different positions as well as different scales (corresponding different frequencies) that distinguish Haar transform from the previous transforms. This capability is also the main advantage of wavelet transform over other orthogonal transforms.

A Haar Transform Example:

The Haar transform coefficients of a $N=4$-point signal $[x[0],x[1],x[2],x[3]]^T=[1,2,3,4]^T$ can be found as

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}=\begin{bmatrix} 5 \\ -2 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

The inverse transform will express the signal as the linear combination of the basis functions:

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & \sqrt{2} & 0 \\ 1 & 1 & -\sqrt{2} & 0 \\ 1 & -1 & 0 & \sqrt{2} \\ 1 & -1 & 0 & -\sqrt{2} \end{bmatrix}\begin{bmatrix} 5 \\ -2 \\ -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$$=\frac{1}{2}[5\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}-2\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \\ 0 \\ 0 \end{bmatrix}-\frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 0 \\ \sqrt{2} \\ -\sqrt{2} \end{bmatrix}]=\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Note that coefficients $X[2]=-1/\sqrt{2}$ and $X[3]=-1/\sqrt{2}$ indicate not only there exist some detailed changes in the signal, but also where in the signal such changes take place (first and second halves). This kind of position information is not available in any other orthogonal transforms.