Wavelets, MRA, Haar wavelet

Lecture notes

1 Introduction

Wavelets are mathematical functions that break up signals into different frequency components, and then study each component with a resolution matched to its scale. Their advantage over traditional Fourier methods is in analyzing signals with discontinuities and sharp spikes. Wavelets were developed independently in the fields of mathematics, quantum physics, electrical engineering, and seismic geology. Interchanges between these fields during the last years have led to many new wavelet applications such as image compression or de-noising data. However, wavelets do not solve all problems and a good alternative are Gabor systems.

The word wavelet is due to Morlet and Grossmann in the early 1980s. They used the French word ondelette, meaning "small wave". Soon it was transferred to English by translating "onde" into "wave", giving "wavelet".

2 Motivation and constrction

In signal analysis it is common to consider signals as square-integrable functions, so here we focus on wavelets in the Hilbert space $L^2(\mathbb{R})$. We are aiming to find some 'simple' functions $\psi_{j,k}$ such that every function $f \in L^2(\mathbb{R})$ has a representation of type

$$f(x) = \sum_{j\in\mathbb{Z}} \sum_{k\in\mathbb{Z}} c_{j,k}\psi_{j,k}(x) \tag{1}$$

for some known coefficients $(c_{j,k})_{j,k\in\mathbb{Z}}$. The motivation comes from approximation theory: The function f might be difficult to work with, but if such a representation exists, then we can hope that the finite partial sums of (1) approximate f well, i.e.

$$f(x) \approx \sum_{|j|\leq M} \sum_{|k|\leq N} c_{j,k}\psi_{j,k}(x). \tag{2}$$

In fact, for wavelets something stronger even holds: under certain conditions, the wavelet expansion converges unconditionally in $L^2(\mathbb{R})$.

We start with a 'nice' function (a wavelet!) ψ and we define a family of functions $\psi_{j,k}$ applying dilation and translation on the original one:

$$\psi_{j,k}(x) := 2^{j/2}\psi(2^j x - k), j, k \in \mathbb{Z}. \tag{3}$$

Observe that $\psi_{0,0}(x) = \psi(x)$ (the original function) and $\psi_{0,k}(x) = \psi(x-k)$ (the translation of the wavelet ψ: the graph of $\psi_{0,k}$ is the graph of ψ, moved by k units to the right). In order to understand the parameter j, we put $k=0$. Then $\psi_{j,0}(x) = 2^{j/2}\psi(2^j x)$, which is the dilated (scaled) version of the original ψ: the larger j is, the more compressed the graph of $\psi_{j,0}$ gets, while negative values of j lead to a less localized version of the graph of ψ.

Our first problem is to determine a function ψ such that every function in $L^2(\mathbb{R})$ has a representation of the form (1). A common constraint on ψ is to induce an orthonormal system, i.e.

$$\langle \psi_{j,k}, \psi_{j',k'} \rangle = \delta_{j-j',k-k'}, \tag{4}$$

which would provide uniqueness for the coefficients $(c_{j,k})_{j,k\in\mathbb{Z}}$ in (1); in fact, we would then have

$$c_{j,k} = \langle f, \psi_{j,k} \rangle. \tag{5}$$

Condition (4) puts too much restriction on ψ: it forces the wavelet system $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ to be an orthonormal basis for $L^2(\mathbb{R})$. It is possible to chose ψ such that (1) holds, even if (4) is not satisfied, but it is then more complicated to calculate the coefficients $(c_{j,k})_{j,k\in\mathbb{Z}}$.

The first example of a function f satisfying (1) was presented by Haar in his PhD thesis in 1910:

$$\psi(x) = \begin{cases} 1 & \text{if } x \in (0, 0.5); \\ -1 & \text{if } x \in (0.5, 1); \\ 0 & \text{elsewhere.} \end{cases}$$

The Haar wavelet satisfies both conditions (1) and (4); it is also compactly supported and is a piecewise polynomial, which properties make it easy to handle. The only disadvantage is that it is discontinuous at $x=0$, $x=0.5$ and $x=1$, which might cause problems in signal transmission: as only a finite number of coefficients $c_{j,k}$ can be transmitted, the receiver will obtain a signal of the form (2), which is discontinuous, even for an infinitely differentiable f. This will have negative effects on transmission of pictures with smooth boundaries, which will appear to be composed of small boxes.

3 MRA

Starting from the '80-ies, there has been an intense research activity aiming at construction of smooth wavelets, with compact support. Around 1987, Mallat and Meyer found a collection of conditions (Multiresolution Analysis, MRA) that lead to construction of wavelets: later in the '90-ies, Daubechies used this result to construct compactly supported wavelet of arbitrary smoothness. Nowadays, an MRA is a standard tool in signal processing. Here is the definition:

Definition 1 *A multiresolution analysis (MRA) consists of a sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$ and a (scaling) function $\phi \in V_0$, satisfying the conditions*

$$
\begin{array}{ll}
(i) & ...V_{-1} \subset V_0 \subset V_1 \subset ... \\
(ii) & \overline{\cup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}),\ \cap_{j\in\mathbb{Z}} V_j = \{0\} \\
(iii) & f \in V_j \Longleftrightarrow f(2\cdot) \in V_{j+1} \\
(iv) & (\forall k \in \mathbb{Z})\, f \in V_0 \Longrightarrow f(\cdot - k) \in V_0 \\
(v) & \{\phi(\cdot - k)\}_{k\in\mathbb{Z}}\ \textit{is an orthonormal basis for } V_0.
\end{array}
$$

By the assumptions in Definition 1, V_1 also has an orthonormal basis: it is $\{2^{1/2}\phi(2\cdot -k)\}_{k\in\mathbb{Z}}$. As $\phi \in V_0 \subset V_1$, there exist coefficients $(c_k)_{k\in\mathbb{Z}}$ such that

$$
\phi(x) = \sum_{k\in\mathbb{Z}} c_k \phi(2x-k), \tag{6}
$$

which gives an exact formulation for the wavelet that generates an orthonormal basis for $L^2(\mathbb{R})$; namely,

$$
\psi(x) = \sum_{k\in\mathbb{Z}} (-1)^k \overline{c_{1-k}} \phi(2x-k). \tag{7}
$$

So, an approximation of a function in $L^2(\mathbb{R})$ can be performed via a function from one of the spaces V_j. The importance of the MRA partly lies in the fact that the conditions guarantee the existance of convenient transforms between the spaces V_j. These transforme, as well as expressions (6) and (7) we shall review on the Haar wavelet example.

4 The Haar wavelet

In terms of piecewise constant functions functions, the Haar wavelet can be written

$$
\psi(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x), \tag{8}
$$

where χ_I is the characteristic function of the set I. In this section, ψ will denote this particular wavelet, and its scaled and translated versions are denoted by $\psi_{j,k}$. The associates scaling function $\phi \equiv \phi_{[0,1)}$ is the characteristic function on $[0,1)$

$$
\phi(x) = \chi_{[0,1)}(x). \tag{9}
$$

Let us consider a continuous function $f:\ [0,1] \rightarrow \mathbb{R}$ and aim to approximate it via piecewise constant functions. The Haar wavelet shall show up in our approximation.

We split the interval $[0,1]$ into 2^k intervals with equal lenght, for some $k = 0, 1, 2, ...$; this means that our intervals are

$$
I_n = \left[n2^{-k}, (n+1)2^{-k}\right), n = 0, 1, ..., 2^k - 1.
$$

The average of f on each interval is

$$
c_n = 2^k \int_{I_n} f(x)dx, \tag{10}
$$

as each interval I_n is of length 2^{-k}.

If k is sufficiently large, that is, the intervals are sufficiently small, the approximationof f (on k-level) is

$$f_k(x) := \sum_{n=0}^{2^k-1} c_n \phi_{I_n}. \tag{11}$$

Comment:

Large values of k give fine approximations; small values give coarse approximations, which are constant on large intervals. We get

Lemma 1 *Assume that $f : [0,1] \to \mathbb{R}$ is continuous. Then, for any $\epsilon > 0$ there exists $K \in \mathbb{N}$ suct that for all $k \geq K$ and all $x \in I_0$*

$$|f(x) - f_k(x)| \leq \epsilon. \tag{12}$$

On k-level, one can describe the k-approximation as

$f_k = f_{k-1} + (f_k - f_{k-1}) = ... = f_0 + \sum_{i=1}^{k} (f_i - f_{i-1})$;

which is the core of the proof of the following theorem

Theorem 1 *Let $f : [0,1] \to \mathbb{R}$ be continuous. Then, for any $k \in \mathbb{N}$, we have*

$$f_k(x) = f_{k-1}(x) + details,$$

and the multiscale representation

$$f_k(x) = c_0 \phi_{[0,1]}(x) + \sum \sum d_{j,n} \psi(2^j x - n), \tag{13}$$

where $d_{j,n} = (c_{j+1,2n} - c_{j+1,2n+1})$ and $c_{j,n} = 2^j \int_{n2^{-j}}^{(n+1)2^{-j}} f(x)dx$.

Observe that the information in the last equation is local: for a given x we only get contributions in the details for the values of j, n such that $2^j x - n \in [0,1]$; for a given j, we only work with the n-values for which $x \in I_n$. This representation also contains frequency information about f: slow/fast oscilations in f lead to nonzero coefficients for small/large values of j.

As for representing the continuous function f, by approximation we get

Theorem 2 *Let $f : [0,1] \to \mathbb{R}$ be a continuous function. Then,*

$$f(x) = \left\langle f, \phi_{[0,1]} \right\rangle \phi_{[0,1]}(x) + \sum_{j=0}^{\infty} \sum_{n=0}^{2^j-1} \left\langle f, \phi_{[0,1]} \right\rangle \psi(2^j x - n). \tag{14}$$

About the singularities:

Walnut's argument shows that, if x_0 is a singularity such that both $f'(x_0-)$ and $f'(x_0+)$ exist, then for lagre j, the coefficients $c_{j,k}$ have different decay depending on the exact location ofx_0; namely,

if $x_0 \in [k2^{-j}, (k+1)2^{-j}]$ then $c_{j,k} \approx 2^{-j/2} \left| f'(x_0-) - f'(x_0+) \right| /4$; while,

if $x_0 \notin [k2^{-j}, (k+1)2^{-j}]$ then $c_{j,k} \approx 2^{-5j/2} \left| f'(x_{j,k}) \right| /4$ for $x_{j,k}$ being the center value of $[k2^{-j}, (k+1)2^{-j}]$.

so, going to higher scales (increasing j), $c_{j,k}$ will be significantly larger around adiscontinuity x_0 and is a good indicatior for a singularity.

References

[1] O. Christensen and Khadija Christensen. Approximation theory: from Taylor polynomials to Wavelets. *Birkhauser*, 2004.

[2] G. Weiss and E.N. Wilson. The Mathematical theory of Wavelets (online ed.). *www.prometheus-us.com.*

[3] I. Daubechies. *Ten lectures on Wavelets.* SIAM, 1992.