Wavelet Basics

(A Beginner's Introduction)

J. S. Marron
Department of Statistics
University of North Carolina

Some references:

Marron, J. S. (1999)"Spectral view of wavelets and nonlinear regression", *Bayesian Inference in Wavelet-Based Models*, Müller, P. and Vidakovic, B. Eds., Lecture Notes in Statistics No. 141, Springer, New York, 19-32.

Strang, G. (1989) Wavelets and dilation equations: a brief introduction, *SIAM Review*, 31, 614-627.

For deeper mathematics, but concisely presented: Chps. 1 and 2 of:

Benedetto, J. J. and Frazier, M. W. (1994) *Wavelets: Mathematics and Applications*, CRC Press, Boca Raton, Florida.

Two Worlds

World 1: "Euclidean vector space",

$$\Re^n = \left\{ \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} : y_1, \ldots, y_n \in \Re \right\}$$

World 2: "(Hilbert) Function Space",

$$L^2 = \left\{ f(x) : \int_0^1 f(x)^2 dx < \infty \right\}$$

Connection: via "digitization"

For equally spaced $0 \le x_1 < \cdots < x_n \le 1$,

Relate $f(x)$ to $\begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}$

Inner Product Structure

World 1:

$$\langle \underset{\rightarrow}{y}, \underset{\rightarrow}{z} \rangle = \left\langle \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \right\rangle = \sum_{i=1}^{n} y_i z_i$$

World 2:

$$\langle f, g \rangle = \int_0^1 f(x) g(x) dx$$

Consequences:

1. "distance" $= \sqrt{\langle a-b, a-b \rangle}$

2. "angle": $a \perp b \quad \Leftrightarrow \quad \langle a, b \rangle = 0$

Connection: Riemann Summation

Linear Bases

$\{y_1, y_2, \ldots\}$ is a "basis" means every member f has a linear representation:

$$f = \sum_i q_i y_i$$

A basis is "orthonormal" when:

$$\langle y_i, y_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

(all orthogonal to each other, with length 1)

Orthonormal Bases

Consequences:

- Compute $q_i = \langle f, y_i \rangle$

- in $\Re^n$, $\underset{\sim}{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$ is the "transform"

- transform is a "rotation" operation
  (lengths and angles preserved)

Example 1: Unit vector basis

$$\underset{\rightarrow}{u}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \Re^n, \qquad i = 1, \ldots, n$$

Notes:

- orthonormal

- for $\underset{\rightarrow}{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, transform has $\boldsymbol{q}_i = y_i$

- "identity rotation"

Example 2: Fourier Basis

Show FourierBasis.ps, with sin's and cos's.

World 1: Discrete Fourier Basis

World 2: Continuous Fourier Basis

Exactly orthonormal in both (takes trigonometry)

Fourier Transform: Rotation that "decomposes into periodicities"

Example 3: Haar Wavelet Basis

Show HaarFullBasis.ps

"Up and Down" step functions, $y_{j,k}$

"doubly indexed" by:

- "scale" j
- "location" k

"dilation form": $y_{j,k}(x) = 2^{-j/2} y\left(2^{-j}x - k\right)$

Exactly orthonormal in both worlds

Dyadic structure, very similar to cascades

Histogram View: successive differences

Show HaarHisto.ps

Example 4: Smoother Wavelet Bases

Daubechies 4: Continuous but "rough"

Show Daub4Basis.ps

Symmlet 8: much smoother, still "local"

Show Symm8Basis.ps

Application 1: Signal Compression

Idea: represent $\underset{\rightarrow}{y}$ by transform $\underset{\rightarrow}{q}$, and hope that "many $q_i \approx 0$"

- "lossless compression", want $q_i = 0$
- "approximate compression", replace q by 0 when "close"

Main Concept:

"Good Compression" $\Leftrightarrow$ more $q_i \approx 0$

Quality of approximation:

Measure by "Energy" in signal:

$$E_{\underline{y}} = \sum_{i=1}^{n} y_i^2 \quad \text{or} \quad E_f = \int_0^1 f(x)^2 dx$$

- lossless compression: $E_{\underline{y}} = E_{\underline{q}}$
  (Parseval Identity)

- Good approximation: $E_{\underline{y}} \approx E_{\underline{q}}$

- Bad approximation: $E_{\underline{y}} >> E_{\underline{q}}$

Approximation Folklore:

Unit vectors: terrible for interesting signals

Fourier basis: good for smooth and periodic

Wavelet bases: allow some jumps

∃ many variations, and ways of "cooking up good bases"

Show ExactRiskEGs.ps and CompressionEG.ps

Application 2: Denoising

Goal: from "data" $\underset{\rightarrow}{y} = \underset{\rightarrow}{s} + \underset{\rightarrow}{n}$

try to recover "signal" $\underset{\rightarrow}{s}$

from "noise" $\underset{\rightarrow}{n}$, (e.g. i.i.d. mean 0)

Transform approach:

- find "rotation" with "good compression of signal"
- zero out small q_i
- invert transform

Denoising Examples

Show WaveDNFourier.eps, StepDNFourier.eps and WaveStepDNHaar.eps

Wave Target:

- Fourier basis: Excellent
- Haar basis: Poor

Step Target:

- Fourier basis: Terrible
- Haar basis: Excellent

Note: driven by signal compression

Fast Computation

of transform: $\mathbf{q}_i = \langle y_i, \mathbf{y}_i \rangle, \quad i = 1, \ldots, n$

1. Naïve implementation: $O(n^2)$ matrix multiplication

2. Fast Fourier Transform: $O(n \log n)$ using trigonometric properties

3. Fast wavelet Transform: $O(n)$ using simple “pyramid algorithm”

Haar Pyramid Algorithm, I

Notation: $\underset{\rightarrow}{1}(n) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \underset{\rightarrow}{0}(n) = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$

"mothers": $\underset{\longrightarrow}{y_{j,k}} = 2^{\frac{-j}{2}} \begin{pmatrix} \underset{\rightarrow}{0}\left(\frac{kn}{2^j}\right) \\ -\underset{\rightarrow}{1}\left(\frac{n}{2^{j+1}}\right) \\ \underset{\rightarrow}{1}\left(\frac{n}{2^{j+1}}\right) \\ \underset{\rightarrow}{0}\left(n - \frac{(k+1)n}{2^j}\right) \end{pmatrix}$

Show HaarFullBasis.ps again

Haar Pyramid Algorithm, II

“fathers”: $$\underrightarrow{j_{j,k}} = 2^{\frac{-j}{2}} \begin{pmatrix} \underrightarrow{0}\left(\frac{kn}{2^j}\right) \\ \underrightarrow{1}\left(\frac{n}{2^j}\right) \\ \underrightarrow{0}\left(n - \frac{(k+1)n}{2^j}\right) \end{pmatrix}$$

Show HaarFathers.ps

Note: father vectors are also a basis (but not orthonormal)

Can mix and match mothers and fathers

Show HaarPartBasis.ps

Haar Pyramid Algorithm, III

Relations across scales:

1. Magnification (dilation):

   j_{j+1} is “half width” of j_j

   y_{j+1} is “halfwidth” of y_j

2. Father $\rightarrow$ Mother, Father

$$\underrightarrow{y_{j,k}} = \frac{1}{\sqrt{2}}\left(\underrightarrow{j_{j+!,2k+1}} - \underrightarrow{j_{j+!,2k}}\right)$$

$$\underrightarrow{j_{j,k}} = \frac{1}{\sqrt{2}}\left(\underrightarrow{j_{j+!,2k+1}} + \underrightarrow{j_{j+!,2k}}\right)$$

Haar Pyramid Algorithm, IV

Apply inner product to get:

$$q_{j,k} = \frac{1}{\sqrt{2}}\left(f_{j+1,2k+1} - f_{j+1,2k}\right)$$

$$f_{j,k} = \frac{1}{\sqrt{2}}\left(f_{j+1,2k+1} + f_{j+1,2k}\right)$$

where

$$f_{j,k} = \left\langle \underrightarrow{j_{j,k}}, \underrightarrow{y} \right\rangle$$

Start with $f_{\log_2(n),k} = y_k$, and iterate up through scales, to get $O(n)$ algorithm

Haar Pyramid Algorithm, V

Overall Structure:

$$f_{j+1,0}, \ldots \quad \leftrightarrow \quad \ldots, f_{j+1,2^{j+1}-1}$$

$$\downarrow \qquad\qquad \downarrow$$

$$lo\ pass\,(avg) \qquad hi\ pass\,(avg)$$

$$\downarrow \qquad\qquad \downarrow$$

$$f_{j,0}, \ldots, f_{j,2^j-1} \qquad q_{j,0}, \ldots, q_{j,2^j-1}$$

$$\leftrightarrow$$

$$\downarrow \qquad \downarrow$$

$$f's \qquad q's$$

Haar Pyramid Algorithm, VI

Notes:

1. each level is "energy preserving":

$$\sum_{k=0}^{2^{j+1}-1} f_{j+1,k}^2 = \sum_{k=0}^{2^j-1} f_{j,k}^2 + \sum_{k=0}^{2^j-1} q_{j,k}^2$$

2. "Energy of constants" passed to f's

3. "Anti-constant energy" passed to q's

Again visit ExactRiskEGs.ps and CompressionEG.ps

4. "Energy issues" are ANOVA style decomposition of sums of squares