CS 3430: SciComp with Py
Assignment 9: Data Analysis with Haar Wavelet Transform

Vladimir Kulyukin
Department of Computer Science
Utah State University

1. Learning Objectives

1. Ordered Haar Wavelet Transform (HWT)
2. Data Analysis
3. Static Methods
4. Technical Writing

2. Introduction

In this assignment, we will use HWT to analyze data. You do not have to write or submit any Py code for this assignment. Instead, your objective is to use the Py implementation of HWT discussed in Lecture 18 to analyze data and briefly summarize the results of your analyses.

3. Ground Frost Depth Analysis with HWT

Recall the ground frost depth problem discussed in Lecture 17 from Ch. 1 of "Wavelets Made Easy" by Y. Nievergelt. A PDF scan of this chapter is attached to the assignment announcement on Canvas. If you read it, you may want to rotate some pages. Here is the problem.

Analyze the following ground frost depth semi-weekly measures, in centimeters, at the Qualchan test station on Hangman Creek from Dec. 1992 to Jan. 1993.

22.0, 27.0, 48.8, 47.5, 47.0, 48.5, 48.0, 47.0, 43.0, 41.0, 41.0, 38.0, 36.0, 47.1, 34.0, 32.0

In **OneDHWTTests.py**, one of the three files in **PySource.zip** attachement to Lecture 18 on Canvas, the data are analyzed in the static method **ex_1_47_p33()** inside the **OneDHWTTests** class:

```
def ex_1_47_p33():
    signal = [22.0, 27.0, 48.8, 47.5, 47.0, 48.5, 48.0, 47.0, 43.0, 41.0, 41.0, 38.0, 36.0, 47.1, 34.0, 32.0]
    OneDHWTTests.testOrdFHWT(signal)
```

Unzip **PySource.zip** into a directory where you will be doing this homework, import **OneDHWTTtests** into the interpreter, and run **ex_1_47_p33**. Here is how it runs in my Py 2.7.6 IDLE:

```
>>> from OneDHWTTests import OneDHWTTests
>>> OneDHWTTests.ex_1_47_p33()
Original signal: [22.0, 27.0, 48.8, 47.5, 47.0, 48.5, 48.0, 47.0, 43.0, 41.0, 41.0, 38.0, 36.0, 47.1, 34.0, 32.0]
Ord FHWT       : [40.49, 1.48, -5.65, 1.74, -11.83, 0.13, 1.25, 4.27, -2.5, 0.65, -0.75, 0.5, 1.0, 1.5, -5.55, 1.0]
```

First, the original values of the signal are printed out. In this case, the values of ground depth frost measurements. Second, the ordered fast haar wavelet transform (Ord FHWT) of the signal is displayed. When you run this example, you will see much greater precision in the transform values. I have rounded each value to 2 digits after the decimal point for simplicity. So, how do we analyze this transform?

Since the measurements are taken semi-weekly, the first eight real values in the original signal are the ground depth frost measurements for the month of January. Specifically, **22.0** and **27.0** are the two measurements for week 1, **48.8** and **47.5** are the two measurements for week 2, etc. Take a look at Lecture 17 for more details.

The first value in the Ord FHWT is **40.49**, which is the average of original signal. In other words, the mean ground frost depth from Dec. 1992 to Jan. 1993 is **40.49cm**. Let us verify that the HWT has given us the correct value.

```
>>> temps = [22.0, 27.0, 48.8, 47.5, 47.0, 48.5, 48.0, 47.0, 43.0, 41.0, 41.0, 38.0, 36.0, 47.1, 34.0, 32.0]
>>> sum(temps)/len(temps)
40.49375
```

The second value is a measurement of change from the first half of the signal to the second half of the signal, or, to put it differently, from December to January. Recall that the real value of the basic Haar wavelet is -2. Thus, the real change from December to January is **(-2)*1.48 = -2.96cm**. The average ground frost depth dropped from December to January by almost **3cm**. Let us verify this.

```
>>> sum(t[:8])/8  ## average frost depth of the first half of the data
41.975
>>> sum(t[8:])/8 ## average frost depth of the second half of the data
39.0125
>>> (41.975-39.0125)/2.0 ## basic wavelet coefficient
1.4812499999999993
>>> (sum(t[:8])/8 – sum(t[8:])/8)/2.0 ## combining all of the above into one line
1.4812499999999993
```

If we add to the average of the first half of data the product of the basic wavelet coefficient and the real magnitude of the basic wavelet, i.e., -2, we should get the average of the second half of data, and we do, indeed:

```
>>> 41.975 + (-2)*1.4812499999993
39.0125000000014
```

Let us dig deeper into the transform values. The next two values, following **1.48**, are **–5.65** and **1.74**. They represent measures of change, i.e., basic wavelet coefficients, from the 1st quarter to the 2nd quarter of the data and from the 3rd quarter to the 4th quarter of the data, respectively. In other words, from the 1st half of December to the 2nd half of December and from the 1st half of January to the 2nd half of January. Verification? Here it is.

```
>>> (sum(t[:4])/4 – sum(t[4:8])/4)/2.0 ## change from 1st half of December to 2nd half of December
-5.649999999999999
>>> (sum(t[8:12])/4 – sum(t[12:])/4)/2.0 ## change from 1st half of January to 2nd half of January
```

```
1.7375000000000007
>>> sum(t[4:8])/4 ## average of 2nd half of December
47.625
>>> sum(t[:4])/4 + (-2)*(-5.65) ## average of 1st half of December plus the wavelet
47.625
>>> sum(t[12:])/4 ## average of 1st half of January
37.275
>>> sum(t[8:12])/4 + (-2)*1.7375000000007 ## average of 1st half of January plus the wavelet
37.2749999999986
>>> sum(t[12:])/4 ## average of the 2nd half of January
37.275
```

The next four values in the transform are **-11.83, 0.13, 1.25, 4.27**. These represent the changes from the 1st week of December to the 2nd week of December (**-11.83**), from the 3rd week of December to the 4th week of December (**0.13**), from the 1st week of January to the 2nd week of January (**1.25**) and from the 3rd week of January to the 4th week of January (**4.27**).

Let us systematically verify these changes one by one starting with the first change, i.e., **-11.83**. Remember I rounded this value to 2 places after the decimal point. The actual value below is **-11.825.**

```
>>> (sum(t[:2])/2 – sum(t[2:4])/2)/2.0 ## change from 1st week of December to 2nd week of December
-11.825
>>> sum(t[:2])/2 + (-2)*(-11.825) ## average of 1st week of December plus the wavelet
48.15
>>> sum(t[2:4])/2 ## average of the 2nd week of December
48.15
```

On to the second change, i.e., **0.13**.

```
>>> (sum(t[4:6])/2 - sum(t[6:8])/2)/2.0
0.125 ## rounded to 0.13 above
>>> sum(t[4:6])/2 + (-2)*0.125
47.5
>>> sum(t[6:8])/2
47.5
```

Here is the verification of the third change, i.e., **1.25.**

```
>>> (sum(t[8:10])/2 - sum(t[10:12])/2)/2.0
1.25
>>> sum(t[8:10])/2 + (-2)*1.25
39.5
>>> sum(t[10:12])/2
39.5
```

Finally, we verify the fourth change, i.e., **4.27**.

```
>>> (sum(t[12:14])/2 - sum(t[14:])/2)/2.0
4.274999999999999
>>> sum(t[12:14])/2 + (-2)*4.27499999999
33.000000000019995
>>> sum(t[14:])/2
33.0
```

The remaining eight coefficients in the transform are **-2.5, 0.65, -0.75, 0.5, 1.0, 1.5, -5.55, 1.0**. They represent measures of change in each of the eight weeks. Let us verify each change coefficient.

```
>>> (t[0]-t[1])/2.0 ## 1st week
-2.5
>>> (t[2]-t[3])/2.0 ## 2nd week
0.6499999999999986 ## rounded to 0.65 above
>>> (t[4]-t[5])/2.0 ## 3rd week
-0.75
>>> (t[6]-t[7])/2.0 ## 4th week
0.5
>>> (t[8]-t[9])/2.0 ## 5th week
1.0
>>> (t[10]-t[11])/2.0 ## 6th week
1.5
>>> (t[12]-t[13])/2.0 ## 7th week
-5.550000000000001 ## rounded to -5.55
>>> (t[14]-t[15])/2.0 ## 8th week
1.0
```

4. Summarizing Your Analysis

If you work as a data analyst, chances are your boss or your client, who asked you to analyze the data, does not know (and, more often than not, does not actually care) what techniques you used to analyze the data. What these people do care about is a comprehensive and brief report of the results. In other words, what do the data tell them? Let us then write a brief technical summary of the above analysis. It may go as follows.

The semi-weekly ground frost depth measurements (in cm) cover the period from December 1992 to January 1993. All measurements are approximate and rounded to 2 places after the decimal point. The average ground frost depth is 40.49cm. From December to January the depth decreased by 3cm. From the 1st half of December to the 2nd half of December the depth increased by 11cm. From the 1st half of January to the 2nd half of January the depth dropped by 3.5cm. From the 1st quarter of December to the 2nd quarter of December, the depth increased by 24cm. From the 3rd quarter of December to the 4th quarter of December, the depth decreased by 0.3cm. From the 1st quarter of January to the 2nd quarter of January, the depth dropped by 2.5cm. From the 3rd quarter of January to the 4th quarter of January, the depth dropped by 8.5cm. Weekly changes were as follows. Week 1 of December: an increase of 5cm;

Week 2 of December: a drop of 1.3cm; Week 3 of December: an increase of 1.5cm. Week 4 of December: an drop of 1cm; Week 1 of January: a drop of 2cm; Week 2 of January: a drop of 3cm; Week 3 of January: an increase of 11cm; Week 4 of January: a drop of 2cm.

5. Doing Your Own Data Analysis

In this section, I will describe two data sets that you will analyze with HWT and write two summaries.

Data Set 1: The semi-weekly measurements below are measurements of river flow, in cubic feet per second, at the US Geological Survey Data Station 1242400 on Hangman Creek, from from December 1992 to January 1993. Analyze them with HWT and write a summary of your analysis similar to the summary in the previous section.

[10.0, 12.0, 12.0, 7.0, 8.0, 9.1, 8.2, 9.4, 16.0, 15.0, 13.0, 11.0, 6.4, 9.0, 19.0, 118.0]

Data Set 2: The file **temps.txt** contains timestamped temperature measurements in degrees Celsius taken in a Langstroth beehive from 12:23 on July 26, 2016 upto 17:54 on August 9, 2016. Although we have never worked with this particular set, we have worked with this format in two previous assignments. I have also massaged the data so that the there are exactly 2048 readings, i.e., an integral power of 2. This will make HWT go smoothly. Analyze the temperature readings with HWT and write up a brief summary that answers the following questions.

1) What is the average temperature?
2) How did the temperature change from the 1st half to the 2nd half of the period?
3) How did the temperature change from the 1st quarter to the 2nd quarter?
4) How did the temperature change from the 3rd quarter to the 4th quarter?

6. What to Submit

Save your summaries in **hwt_analysis.txt** and submit it through Canvas. You do not have to submit any Py source.

Enjoy your analysis!