Statistical Engineering Division Workshop

Spectral Methods for Data Analysis

Mark S. Levenson

June 24, 1997
9-12 a.m.
Administration Building
Lecture Room A

Mark S. Levenson
Bldg. 820 Rm. 353
National Institute of Standards and Technology
Gaithersburg, MD 20899
mark.levenson@nist.gov

Outline

1: Motivating Examples

2: Mathematical and Statistical Background

3: Fourier Analysis

4: Wavelet Analysis

5: Principal Component Analysis

References

Linear Algebra and Its Applications, G. Strang.

Numerical Recipes in C, Second Edition, W. Press, et al.

"Wavelet transforms versus Fourier transforms," *Bulletin American Mathematical Society*, G. Strang, (1993).

"Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, D. L. Donoho and I. M. Johnstone, (1994).

Applied Multivariate Statistical Analysis, R. Johnson and D. Wichern.

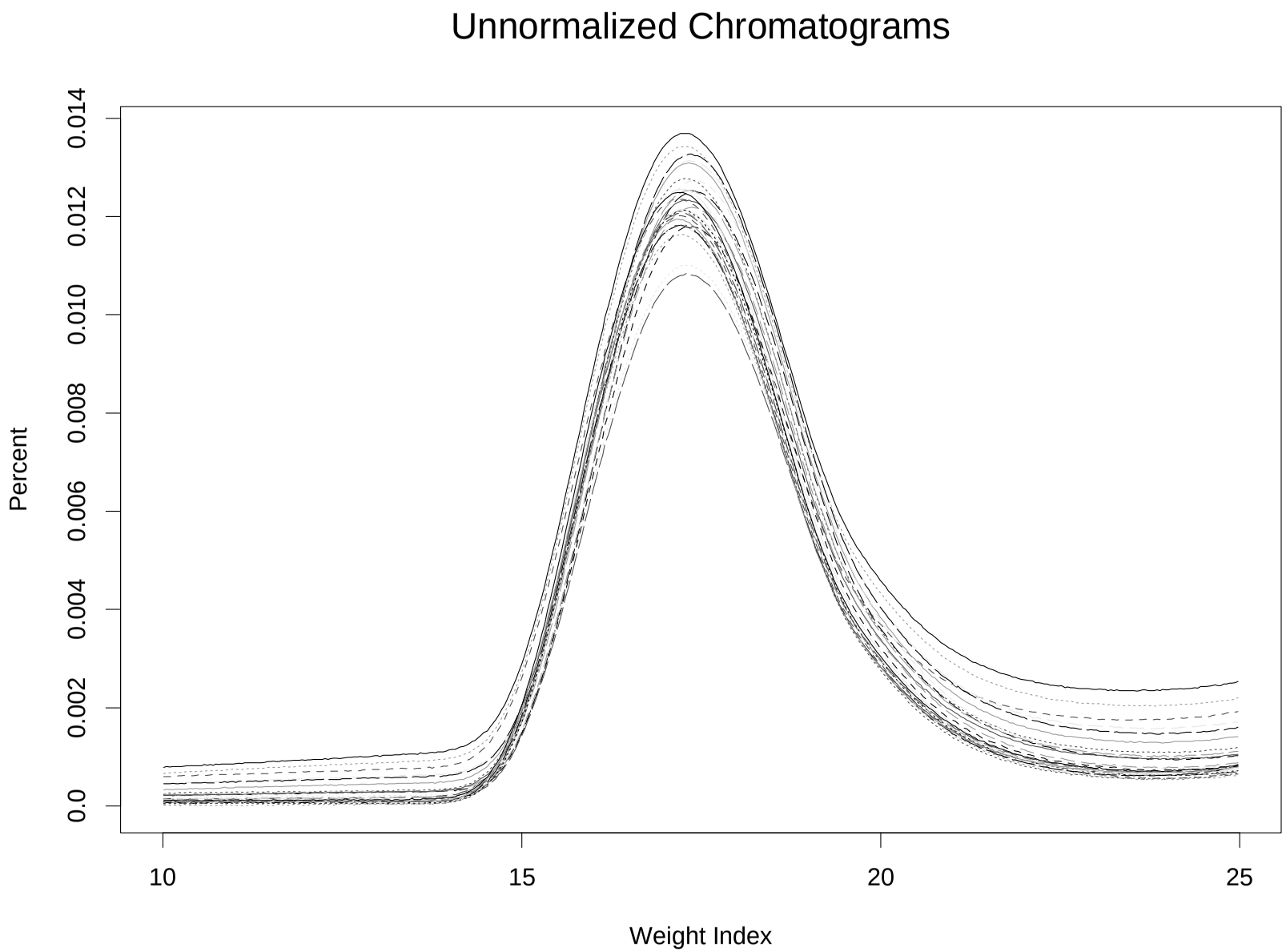
1: Motivating Examples

a: Chromatography Spectra

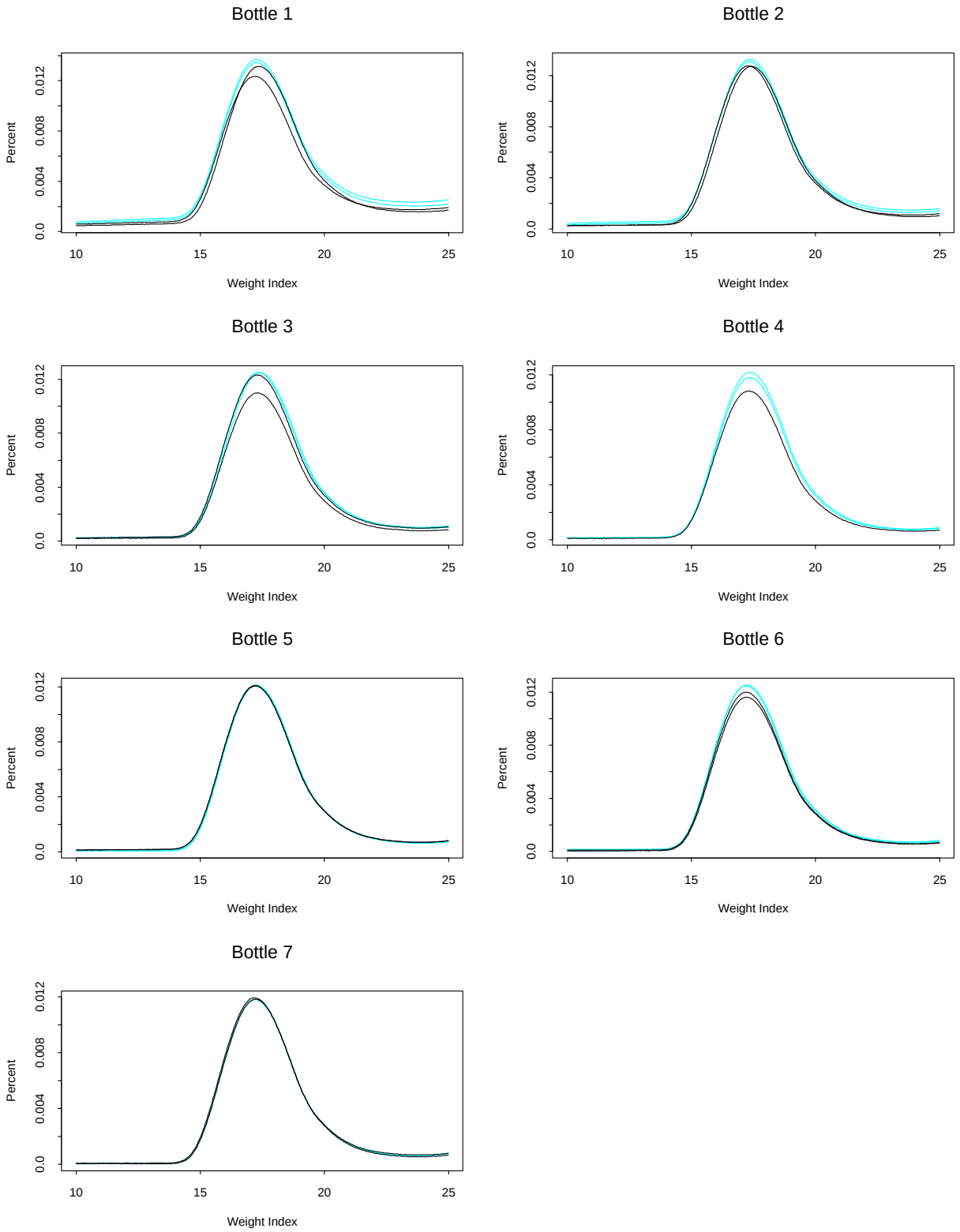
b: Turbulence Study

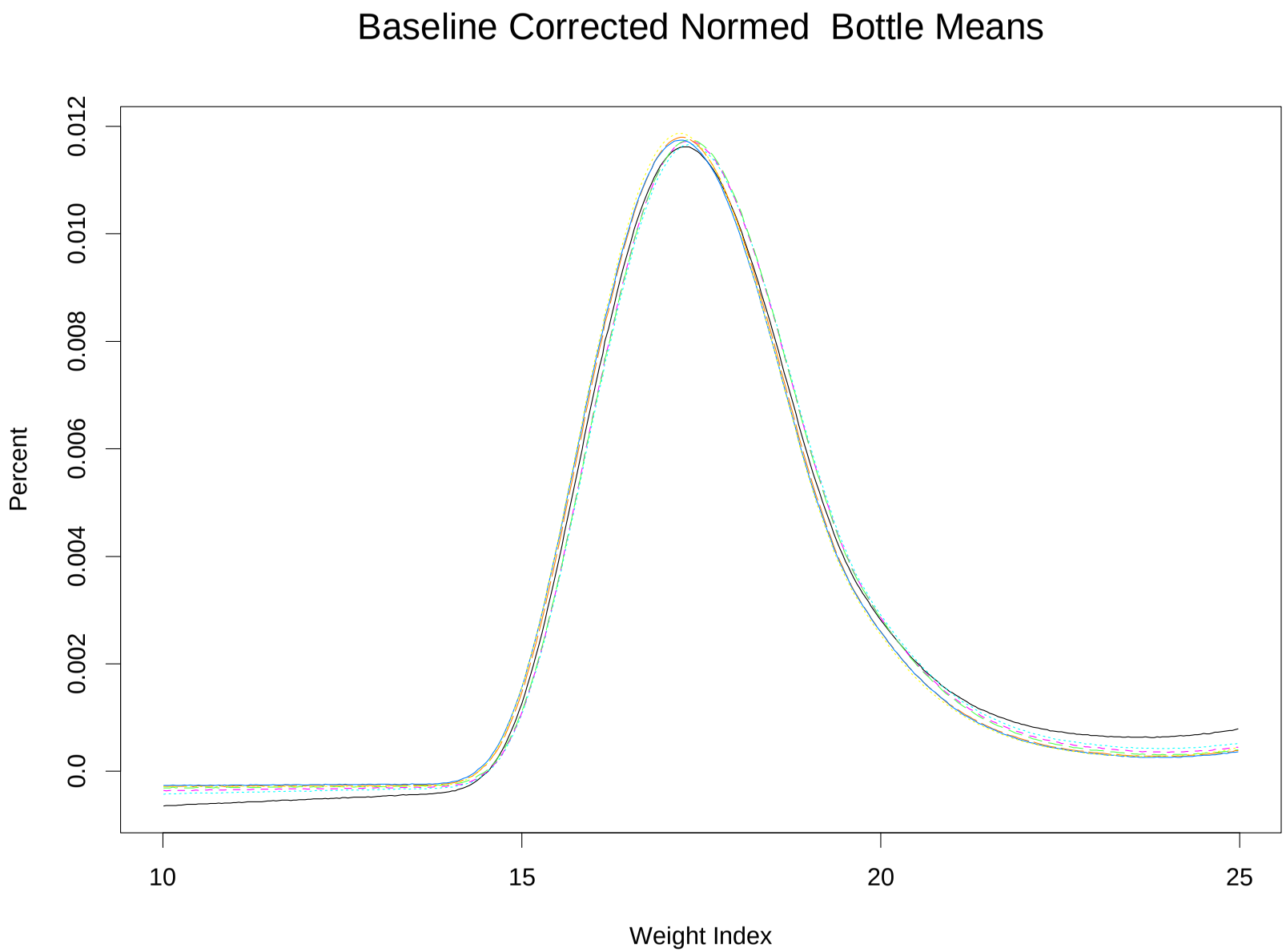
c: Random Bits

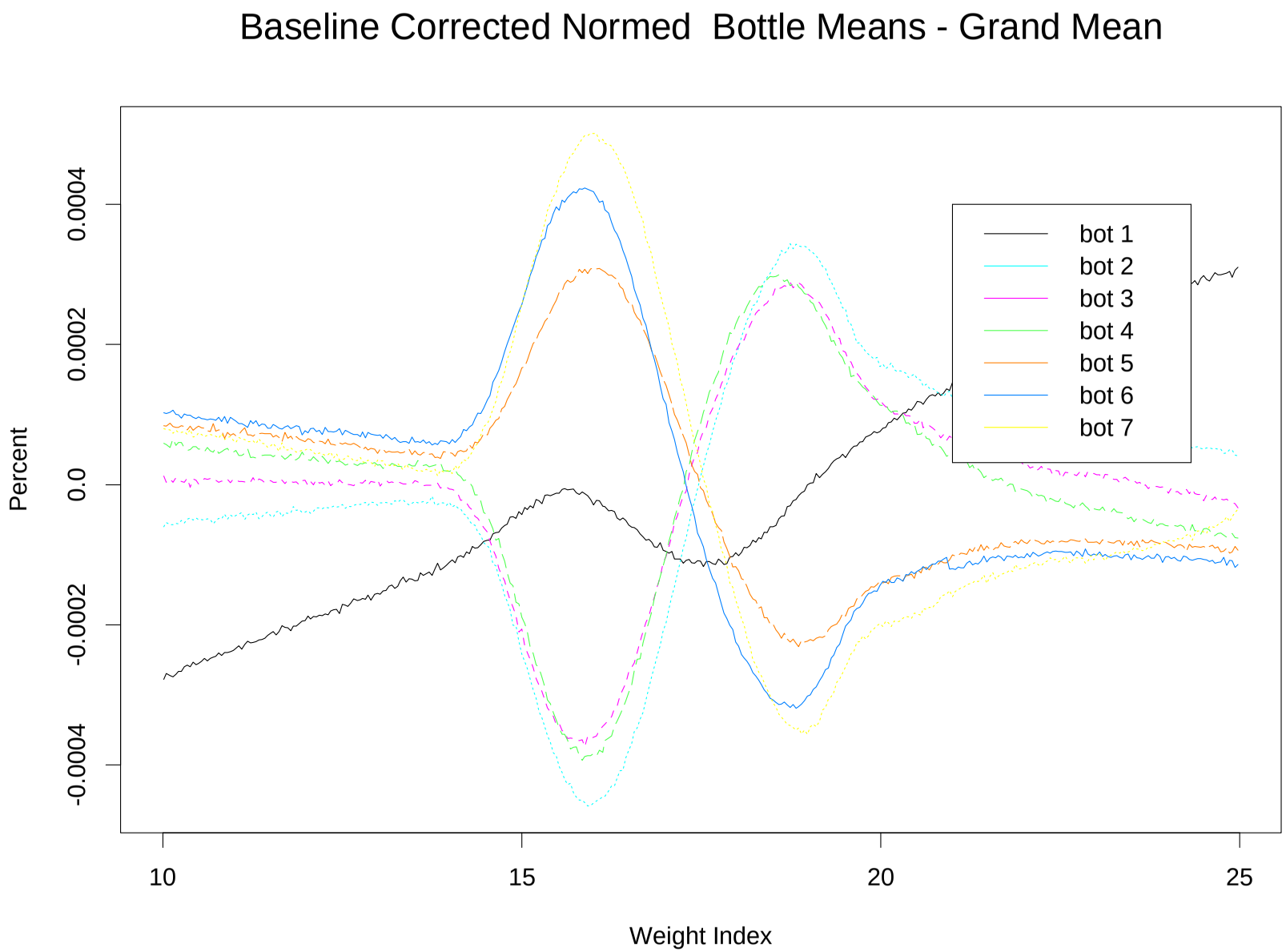
1a: Chromatography Spectra

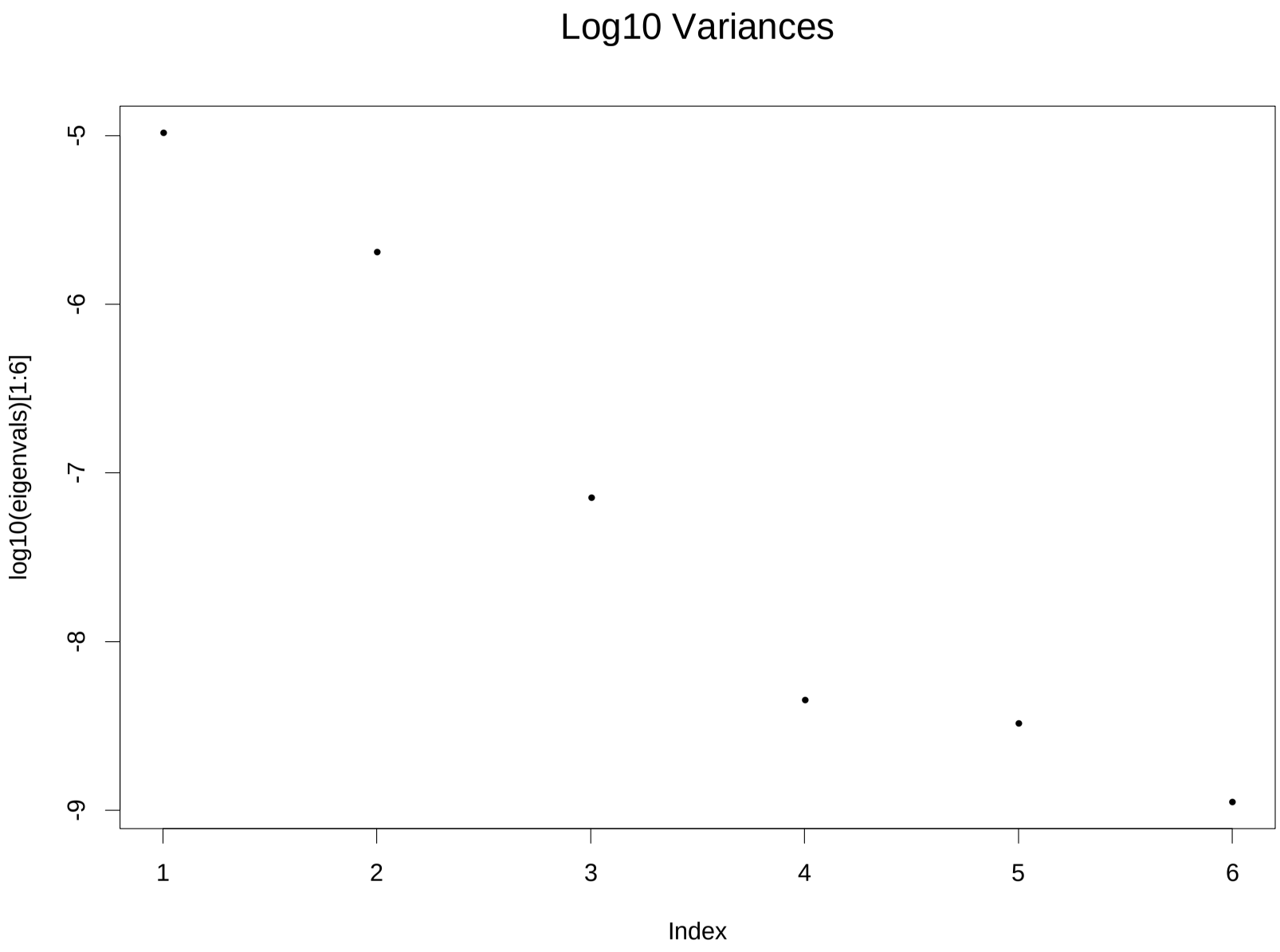

Unnormalized Chromatograms
Percent
0.0
0.002
0.004
0.006
0.008
0.010
0.012
0.014
10
15
20
25
Weight Index

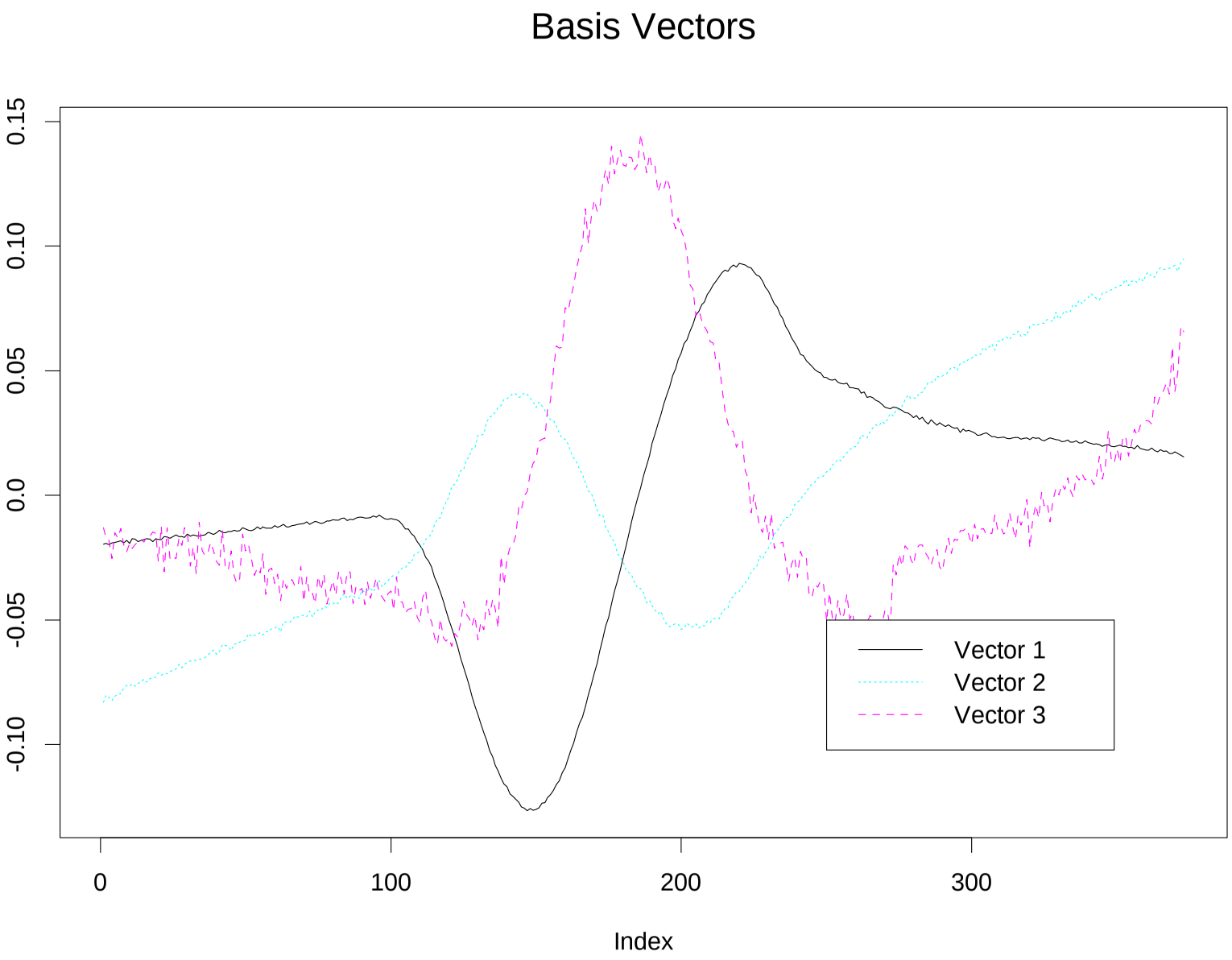
Unnormalized Chromatograms (Color=Vial)

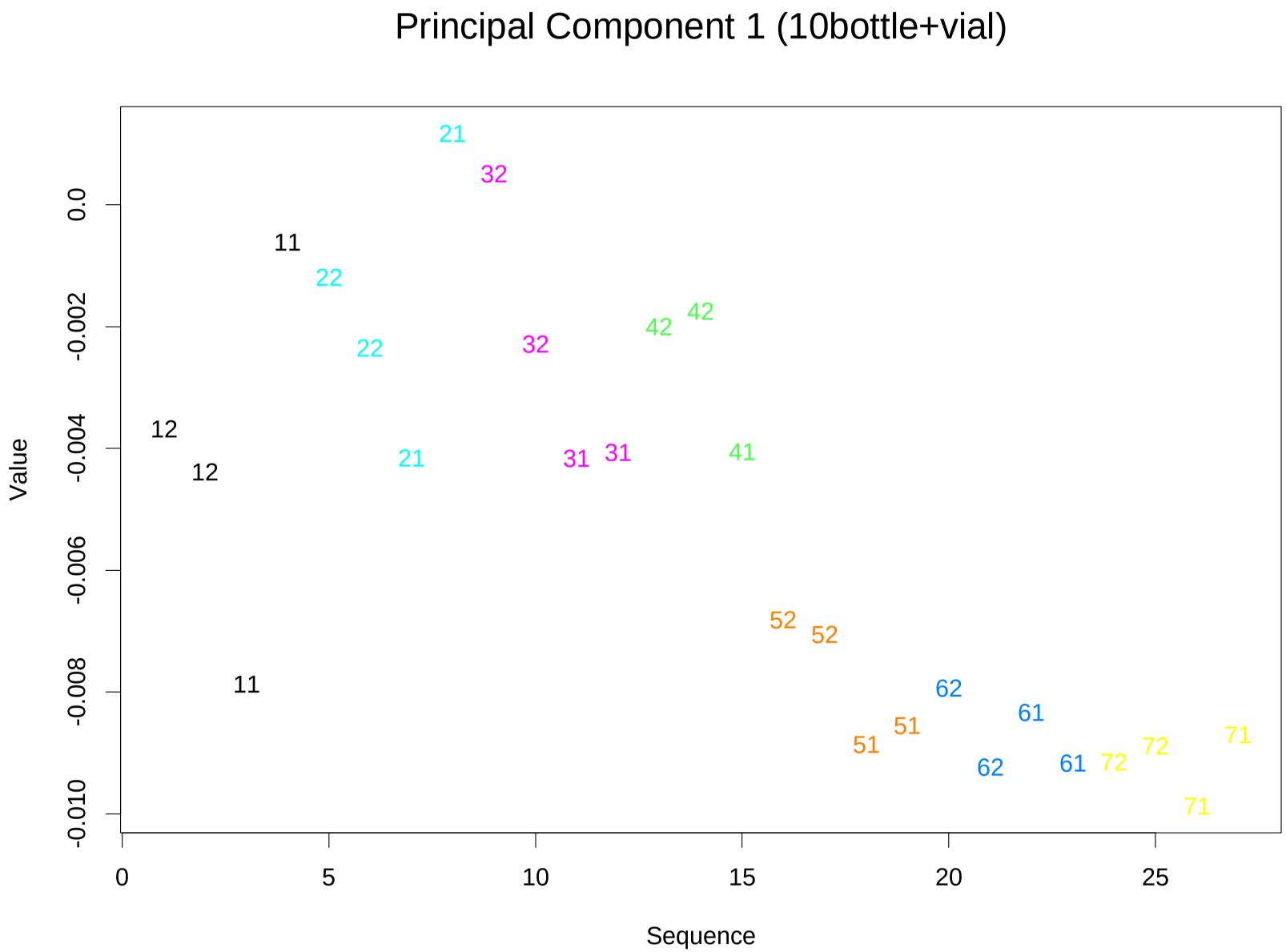


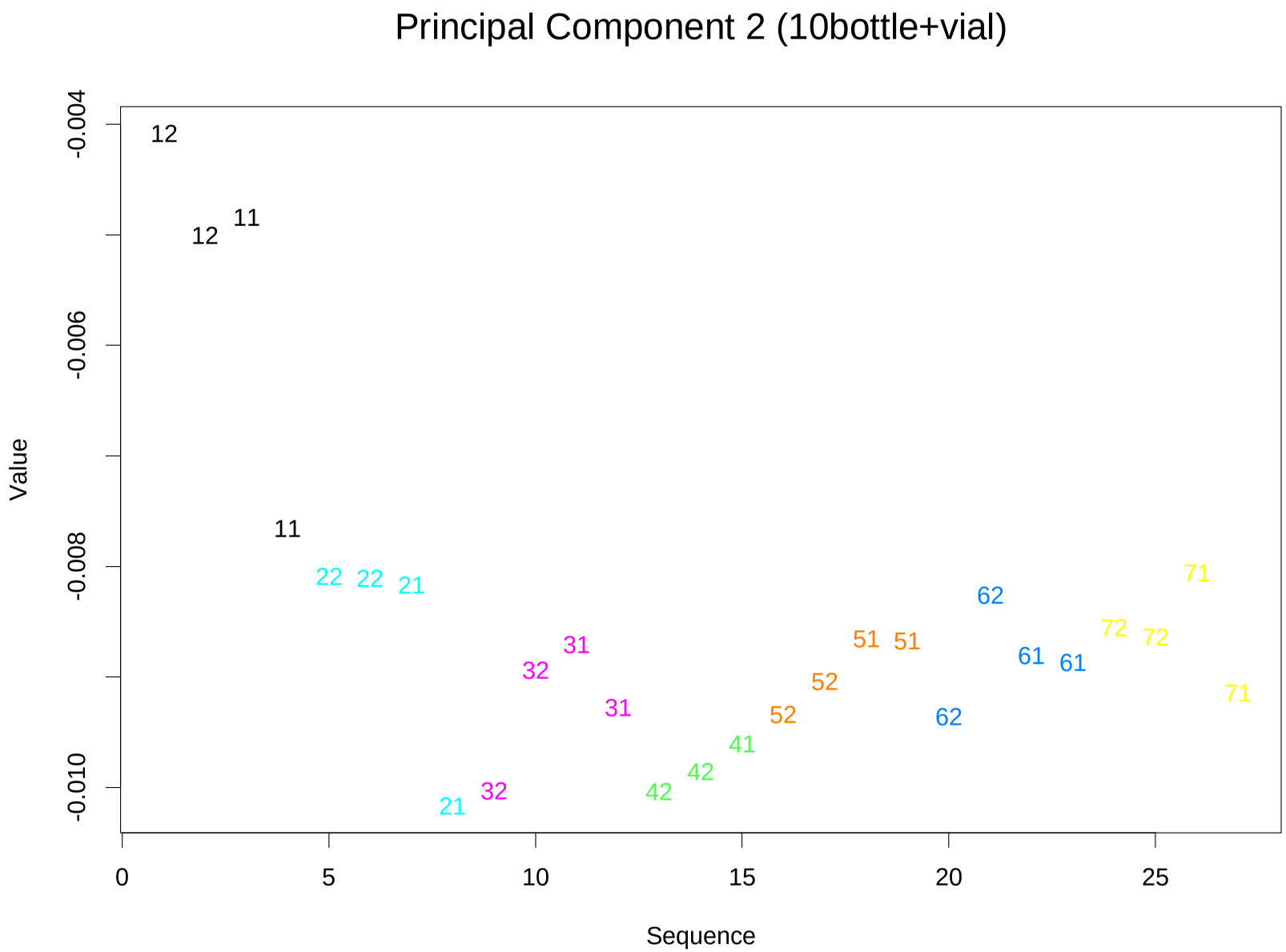

Baseline Corrected Normed Bottle Means
Percent
0.0
0.002
0.004
0.006
0.008
0.010
0.012
10
15
20
25
Weight Index

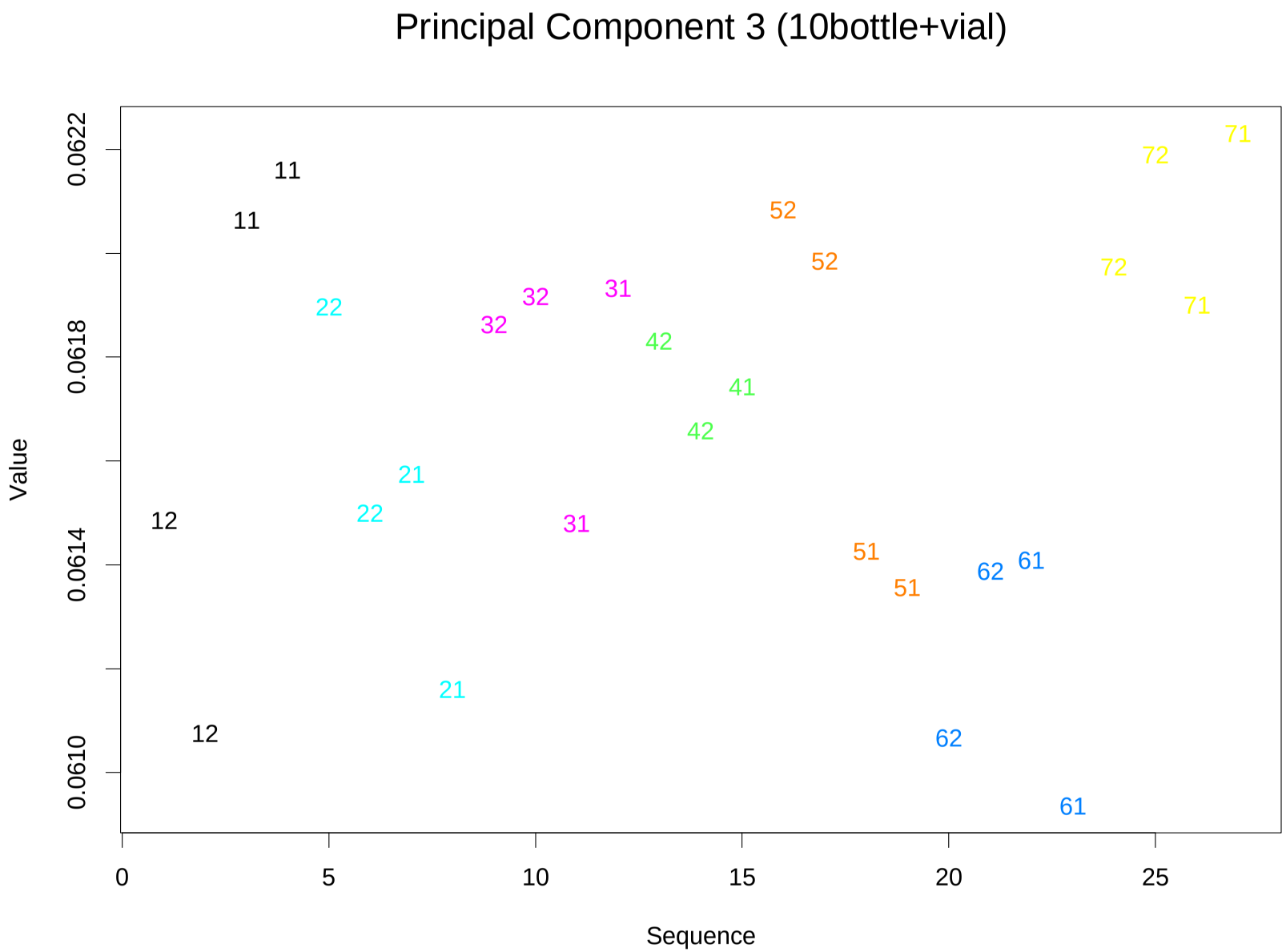

Baseline Corrected Normed Bottle Means - Grand Mean
Percent
0.0004
0.0002
0.0
-0.0002
-0.0004
10
15
20
25
Weight Index
bot 1
bot 2
bot 3
bot 4
bot 5
bot 6
bot 7


Log10 Variances
log10(eigenvals)[1:6]
-5
-6
-7
-8
-9
1
2
3
4
5
6
Index

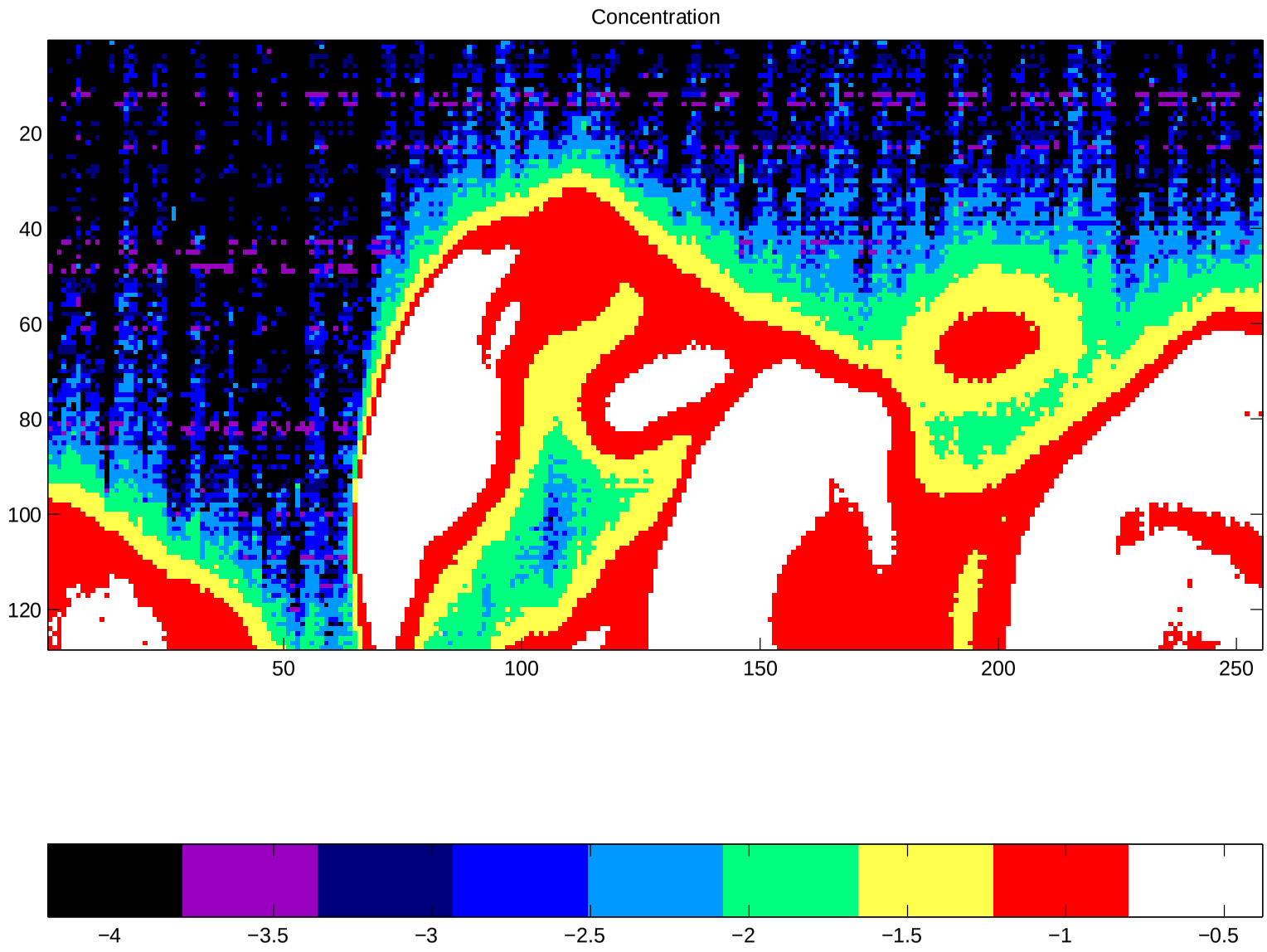

Basis Vectors
0.15
0.10
0.05
0.0
-0.05
-0.10
0
100
200
300
Index
Vector 1
Vector 2
Vector 3

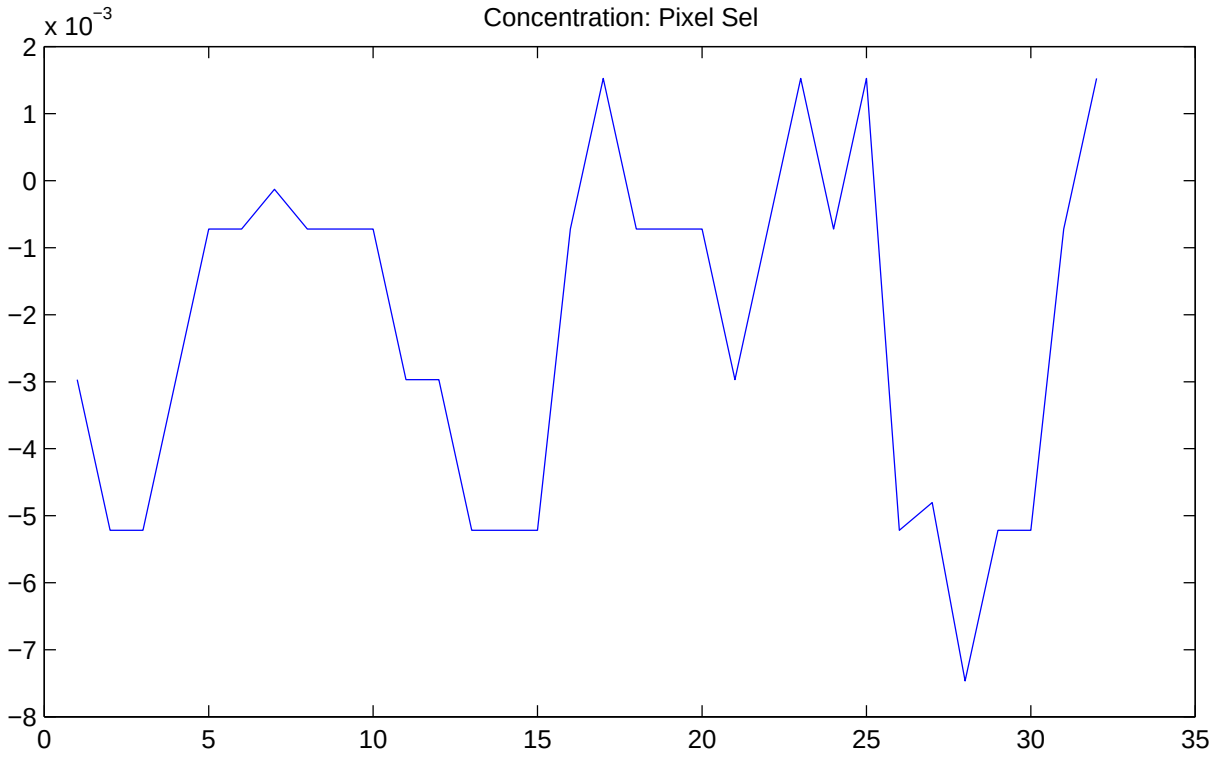

Principal Component 1 (10bottle+vial)
Value
0.0
-0.002
-0.004
-0.006
-0.008
-0.010
0
5
10
15
20
25
Sequence
21
32
11
22
42
42
22
32
12
21
31
31
41
12
52
52
11
62
61
51
51
62
61
72
72
71
71

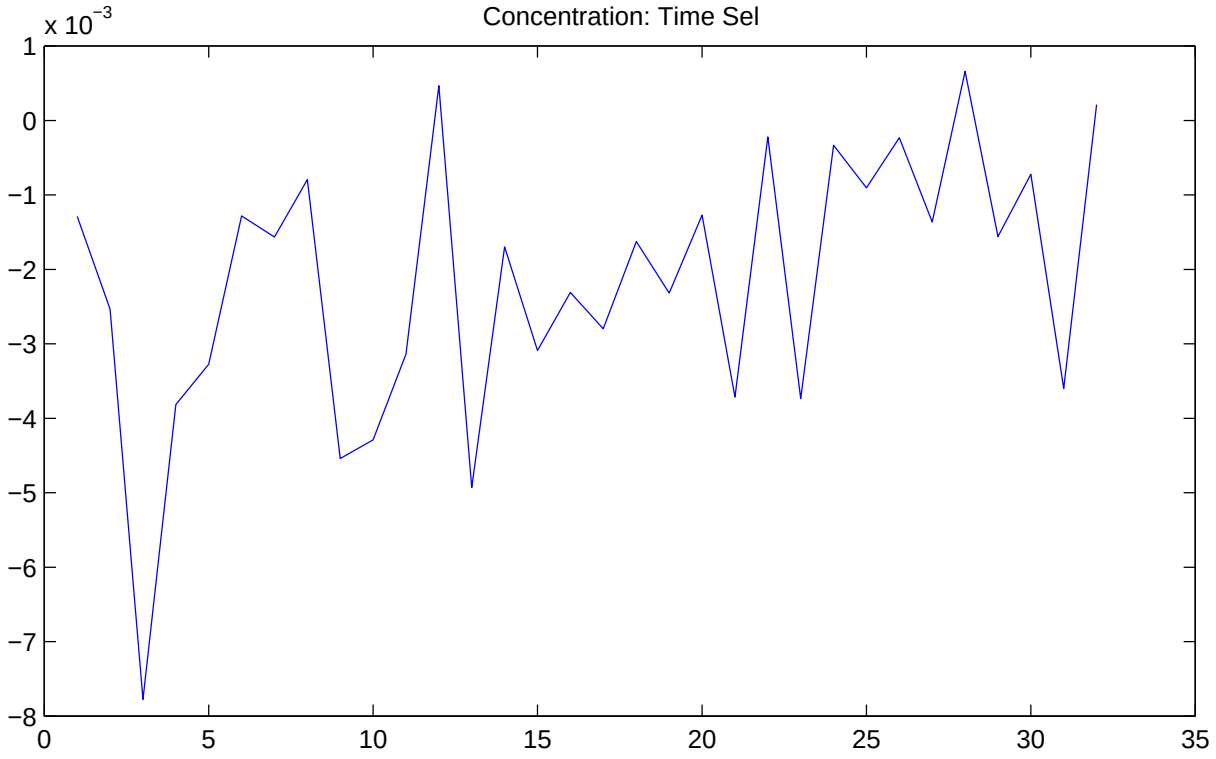

Principal Component 2 (10bottle+vial)
Value
-0.004
-0.006
-0.008
-0.010
Sequence
0
5
10
15
20
25
12
12
11
11
22
22
21
21
32
32
31
31
42
42
41
52
52
51
51
62
62
61
61
72
72
71
71

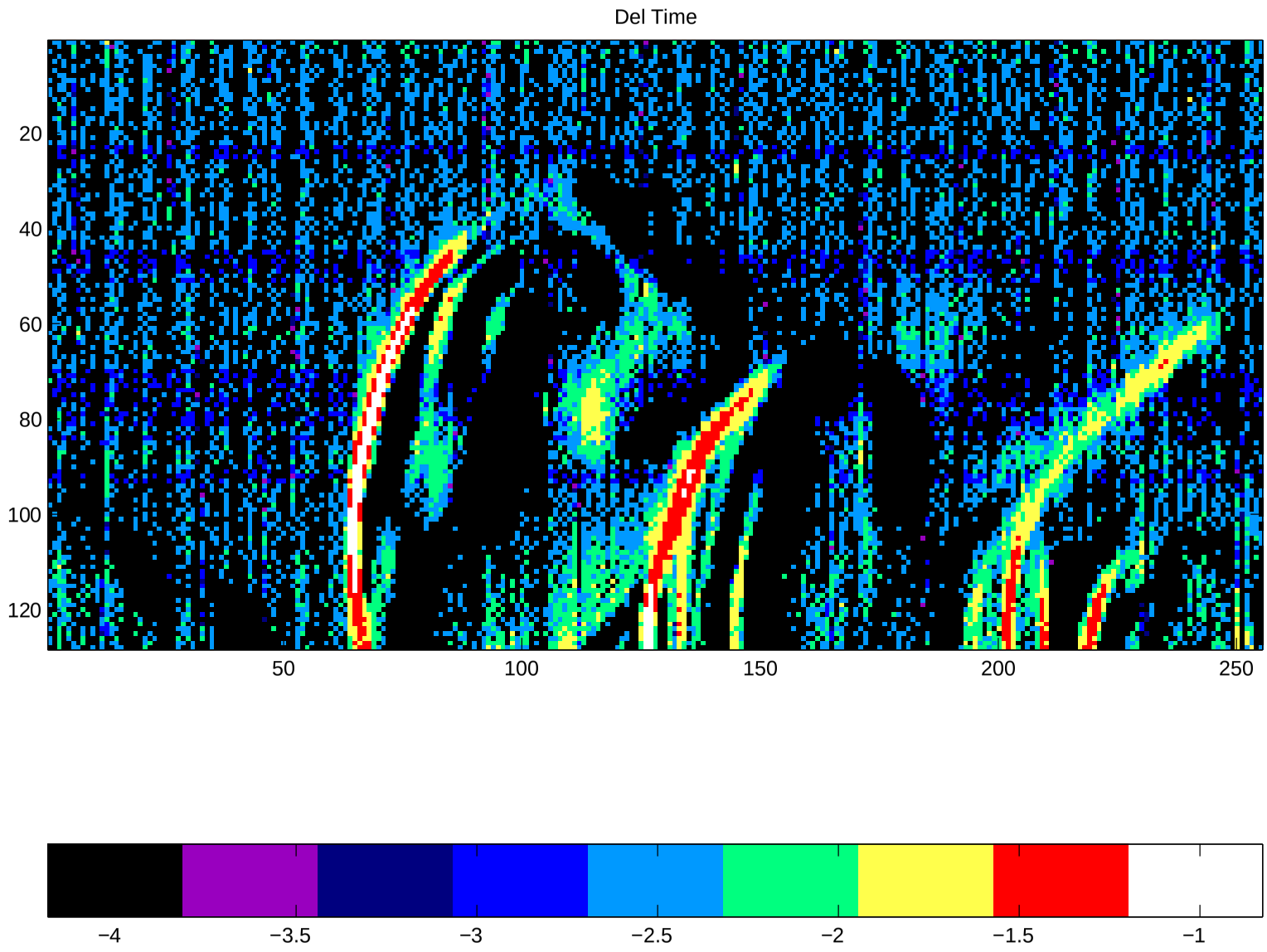

Principal Component 3 (10bottle+vial)
Value
0.0610
0.0614
0.0618
0.0622
Sequence
0
5
10
15
20
25
11
11
12
12
22
22
21
21
32
32
31
31
42
42
41
52
52
51
51
62
62
61
61
72
72
71
71

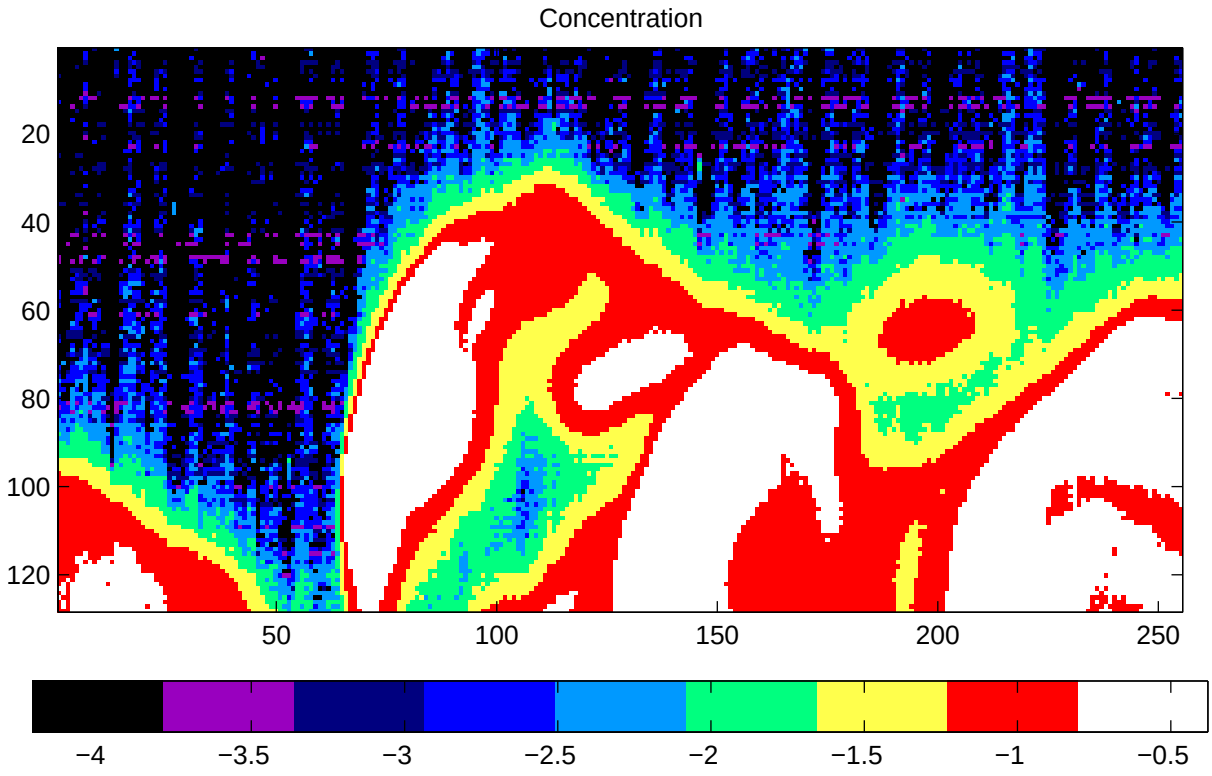
1b: Turbulence Study

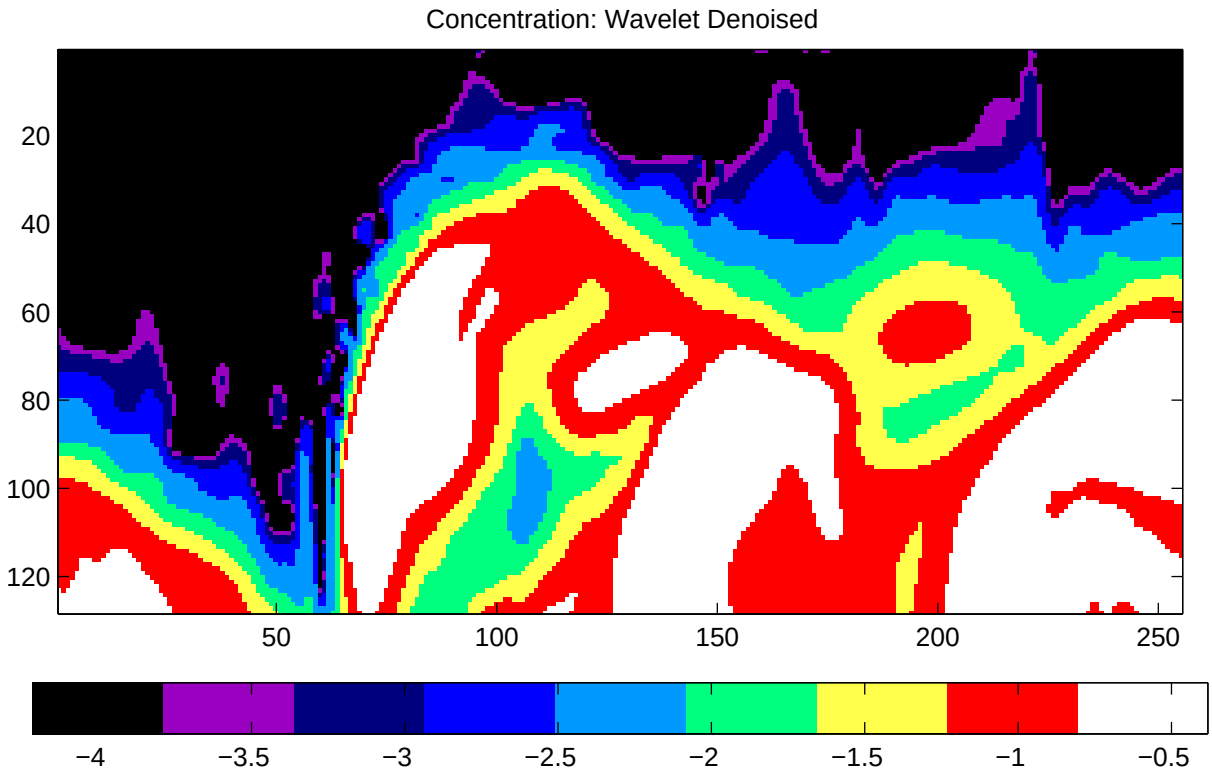

Concentration
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5
−1
−0.5

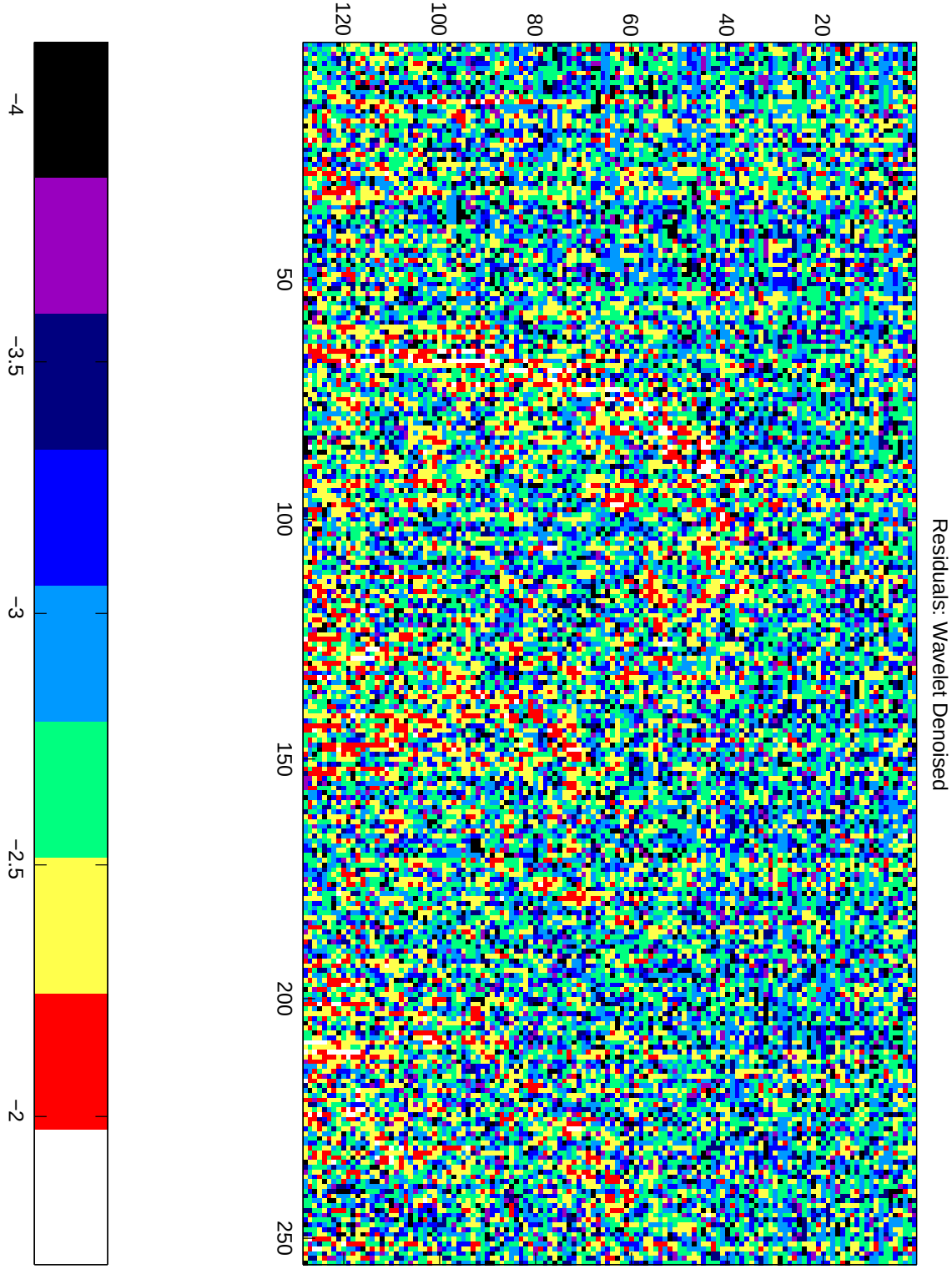

Concentration: Pixel Sel
x 10^-3
2
1
0
−1
−2
−3
−4
−5
−6
−7
−8
0
5
10
15
20
25
30
35

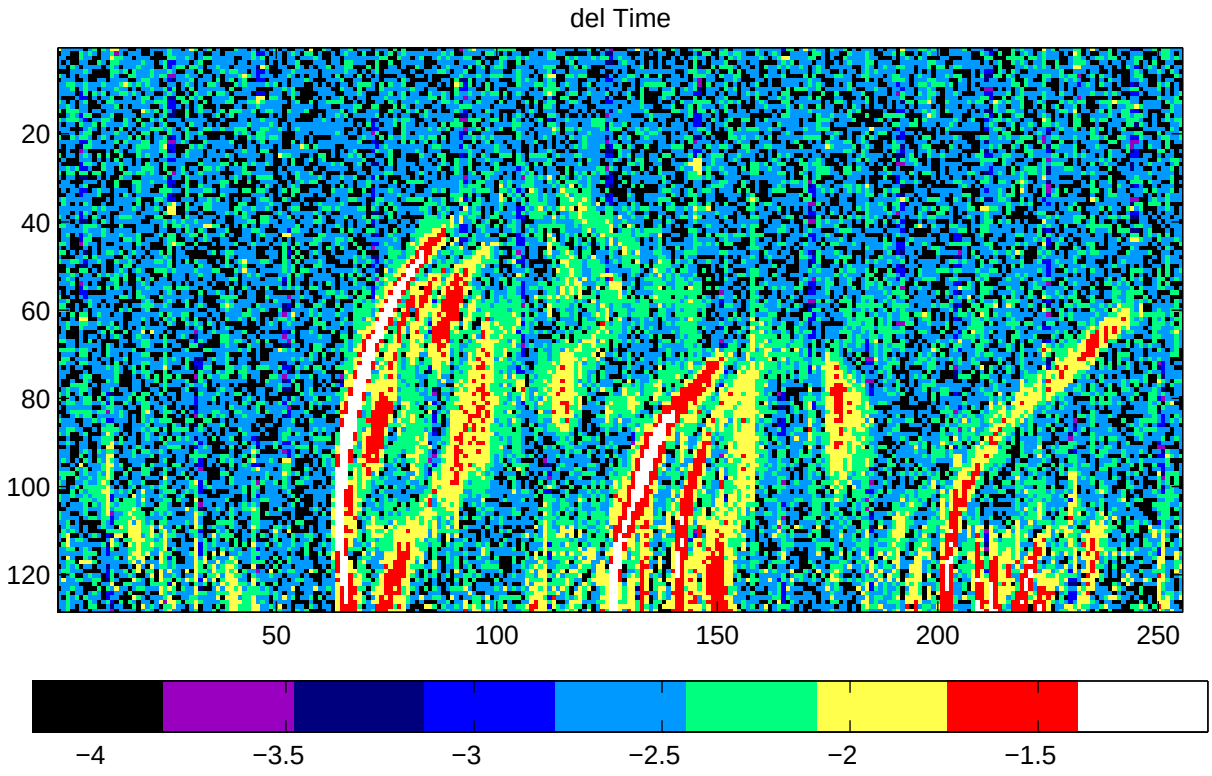

Concentration: Time Sel
x 10^-3
1
0
−1
−2
−3
−4
−5
−6
−7
−8
0
5
10
15
20
25
30
35

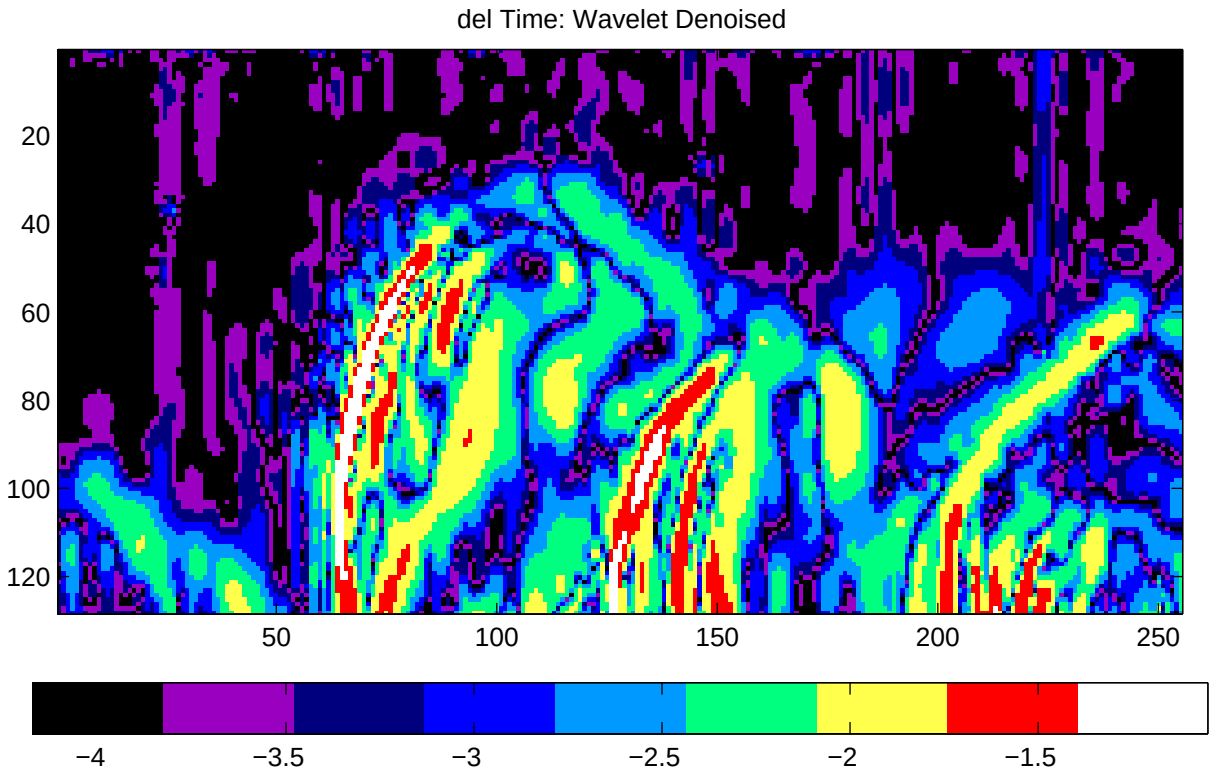

Del Time
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5
−1

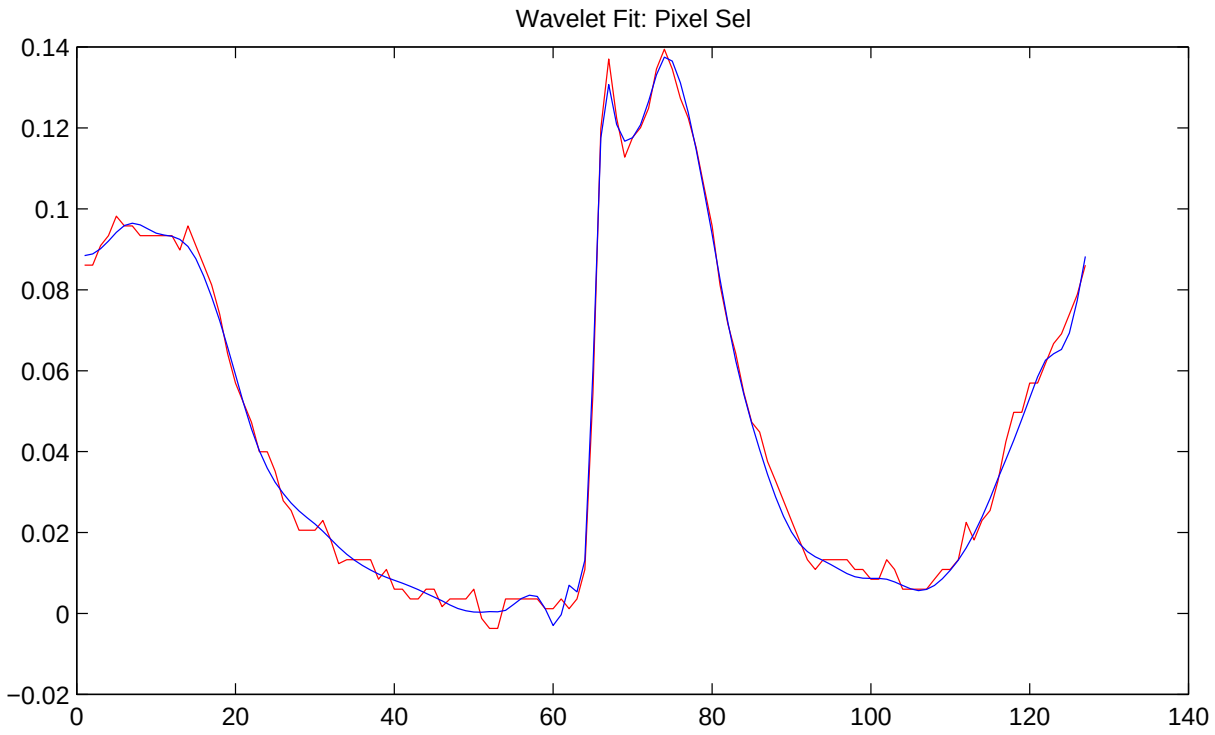

Concentration
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5
−1
−0.5

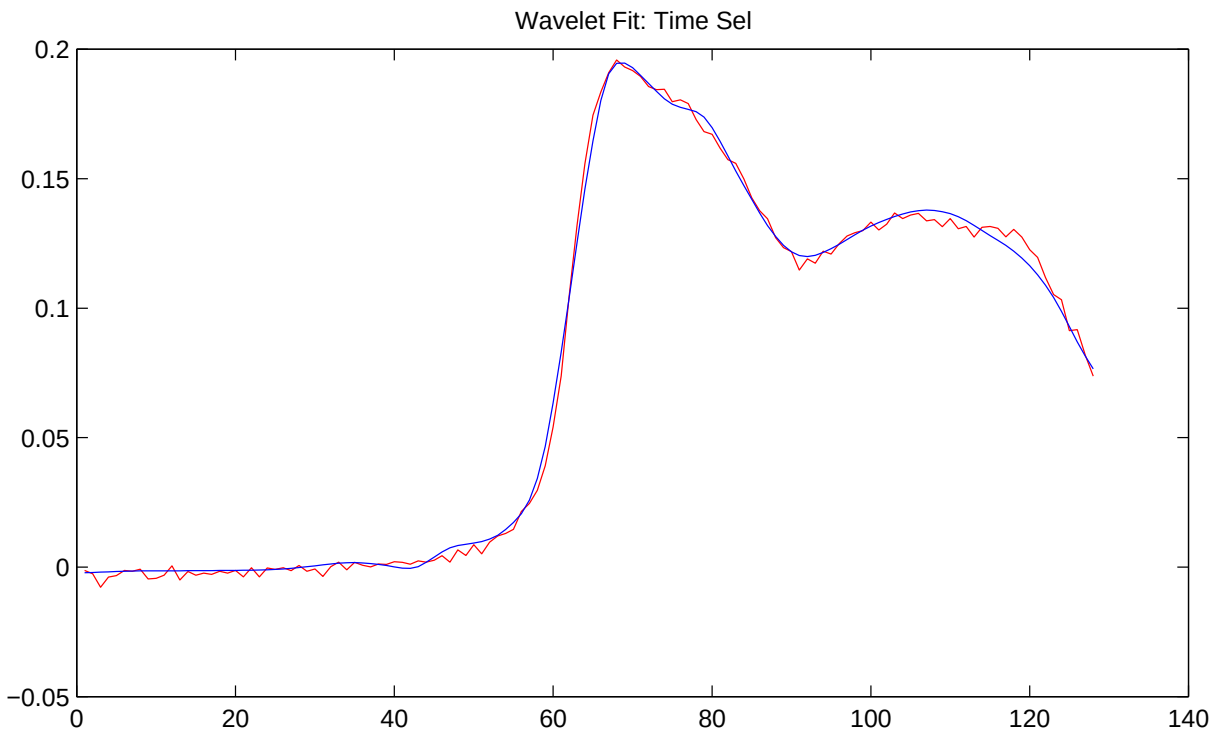

Concentration: Wavelet Denoised
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5
−1
−0.5

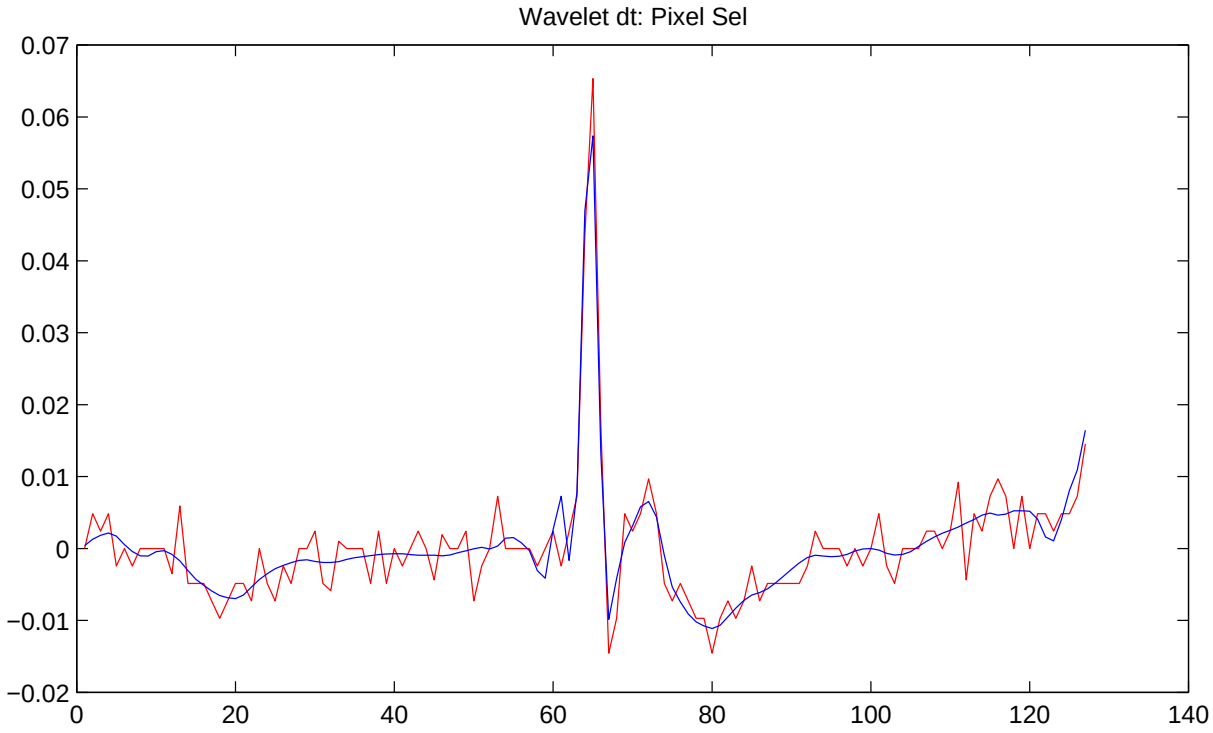

Residuals: Wavelet Denoised
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2

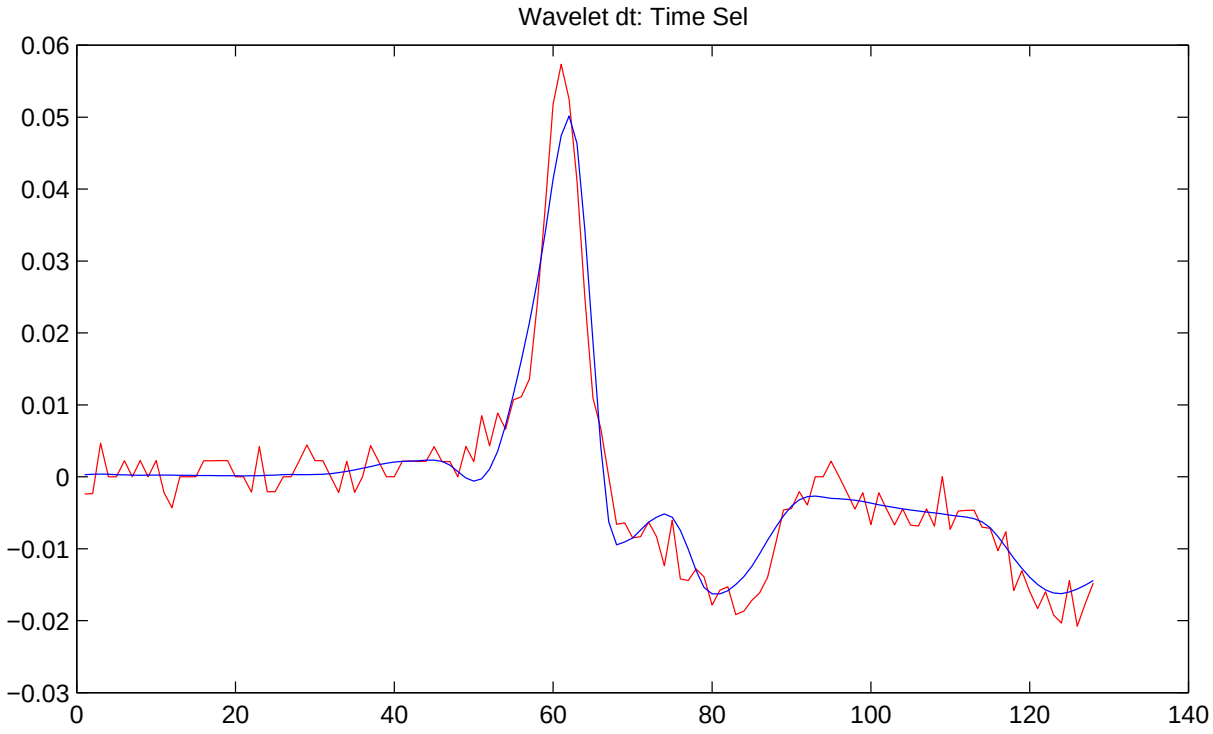

del Time
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5


del Time: Wavelet Denoised
20
40
60
80
100
120
50
100
150
200
250
−4
−3.5
−3
−2.5
−2
−1.5


Wavelet Fit: Pixel Sel
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
−0.02
0
20
40
60
80
100
120
140


Wavelet Fit: Time Sel
0.2
0.15
0.1
0.05
0
−0.05
0
20
40
60
80
100
120
140

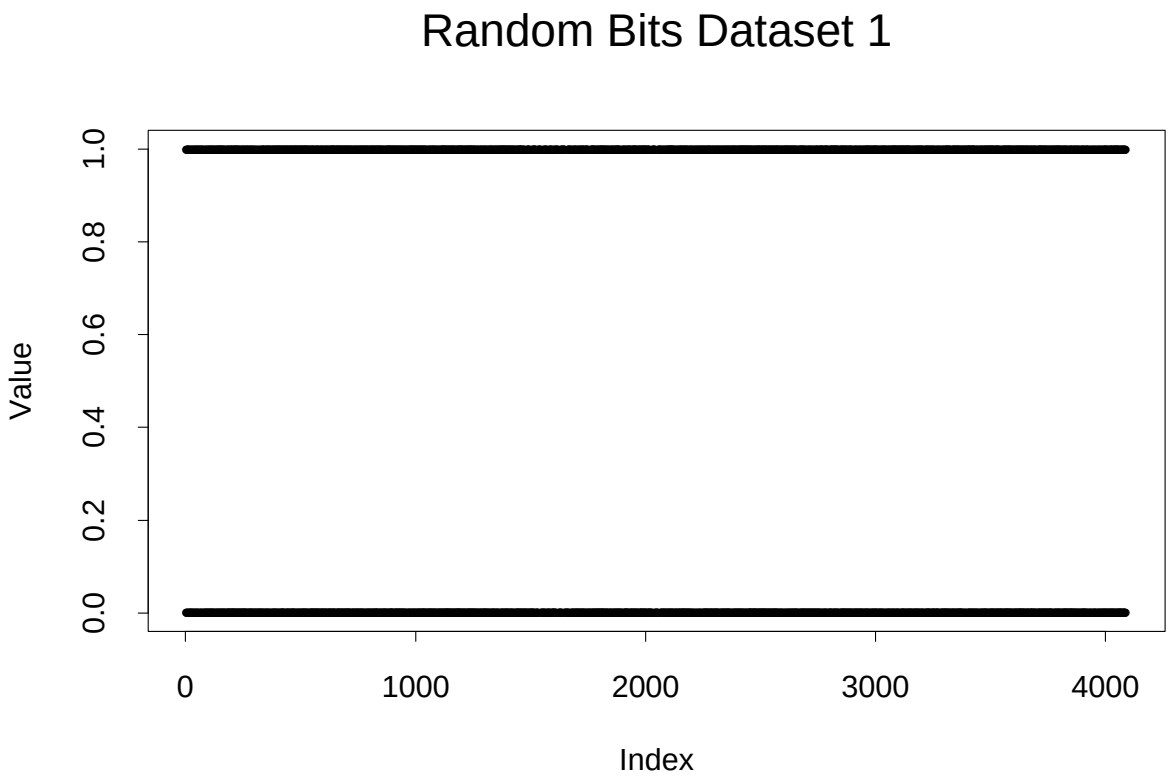

Wavelet dt: Pixel Sel
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
−0.02
0
20
40
60
80
100
120
140

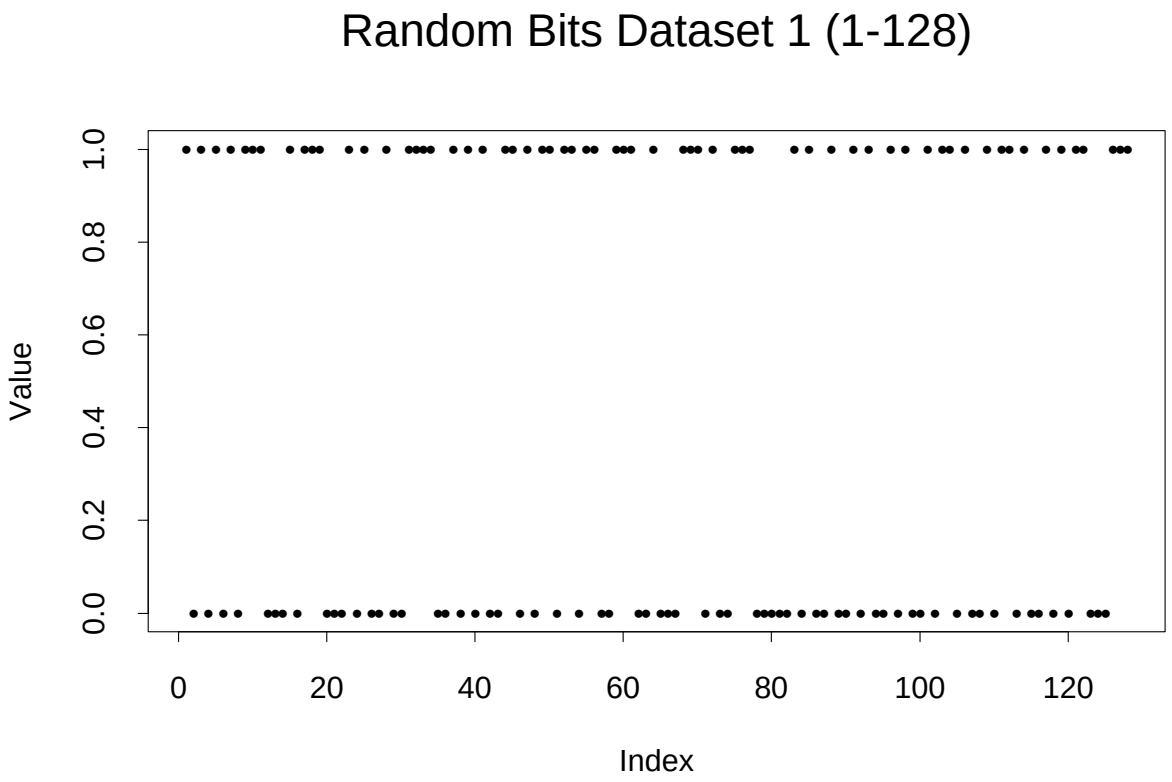

Wavelet dt: Time Sel
0.06
0.05
0.04
0.03
0.02
0.01
0
−0.01
−0.02
−0.03
0
20
40
60
80
100
120
140

1c: Random Bits

Dataset 1 Random Bits

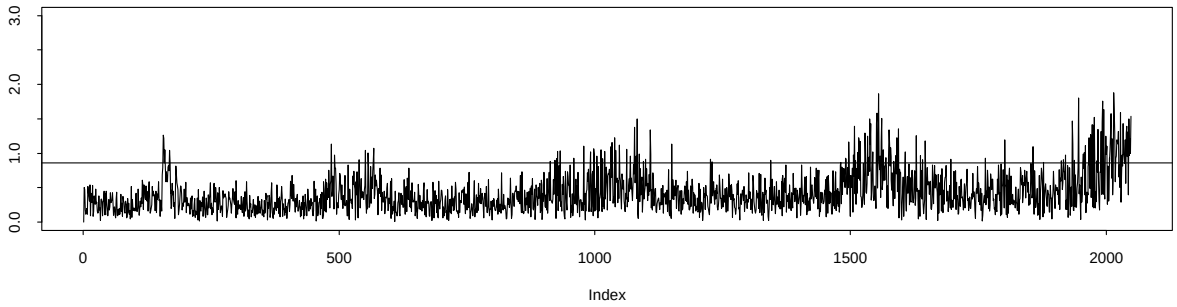
```
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø 1 Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø 0 1 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1
Ø Ø 0 Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø
Ø Ø Ø 1 Ø Ø Ø Ø Ø 0 Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø
Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø 1 Ø 1
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø 0 Ø 0 Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø 1 0 Ø Ø Ø Ø Ø Ø 1 Ø Ø 1 Ø 1 Ø Ø Ø Ø Ø
Ø Ø Ø Ø 1 Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø 1 Ø 1 Ø Ø Ø 0 Ø Ø 1 0 Ø Ø Ø 0 Ø Ø
Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 0 Ø 1 Ø 1 Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø
0 Ø Ø Ø Ø Ø Ø 1 0 Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø 1 0 Ø Ø Ø 0 1 0 Ø Ø Ø Ø
Ø Ø 1 0 Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø 0 Ø Ø Ø 0 Ø Ø Ø Ø 0 Ø Ø Ø Ø 1 0
1 Ø 0 Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 0 Ø Ø Ø 1 Ø 0 Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø 0
Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø
Ø 1 Ø 0 Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 0 Ø 1 Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0
Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø 1 0 Ø Ø Ø 1 Ø Ø 1 0 1 Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø 0 1
0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø 1 Ø Ø Ø Ø
Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø
Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø
Ø 0 Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø 1 0 Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø
0 1 Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 0 Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø
1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø
Ø 1 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø 0 1 Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø
1 Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
0 1 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0
Ø Ø Ø Ø Ø 0 Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø 0
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø 1 Ø Ø 1
Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø
Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 0 Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 0 Ø Ø Ø 0 Ø Ø
Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 0 Ø 0 Ø Ø Ø 0 Ø
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø 0 1 Ø 0 Ø Ø Ø 0 Ø Ø Ø 0 1 Ø Ø Ø Ø Ø
Ø Ø Ø 1 Ø 0 1 Ø Ø 0 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø 1 0 Ø 0 1 Ø
Ø Ø 1 Ø 0 Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 0 Ø 0 Ø
0 Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 1 0 Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø
Ø Ø 1 Ø Ø Ø Ø 1 Ø 0 Ø Ø Ø Ø 0 1 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø 0 Ø
Ø Ø Ø Ø Ø Ø Ø Ø 1 0 0 1 0 Ø 0 1 Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 1 0 Ø
Ø Ø Ø Ø Ø Ø Ø 0 Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0
Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 1 Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 1 Ø Ø Ø
Ø 0 Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø
Ø Ø Ø Ø Ø Ø 1 0 1 Ø Ø Ø Ø Ø Ø Ø Ø Ø Ø 0 Ø 1 Ø Ø Ø Ø Ø 0 Ø Ø 1 Ø
Ø Ø Ø 0 Ø Ø 1 0 1 Ø Ø Ø Ø 0 Ø 0 1 Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø Ø 1 Ø Ø Ø
Ø 0 Ø Ø Ø 0 Ø Ø 0 0 0 Ø Ø Ø Ø Ø Ø 1 1 Ø Ø Ø Ø 1 Ø Ø Ø Ø Ø 0 Ø 0
Ø Ø Ø 0 0 Ø Ø 0 1 Ø Ø Ø 1 Ø 1 Ø Ø Ø 1 0 1 Ø Ø Ø 0 Ø 0 Ø Ø Ø Ø Ø
```


Random Bits Dataset 1
Value
0.0
0.2
0.4
0.6
0.8
1.0
0
1000
2000
3000
4000
Index

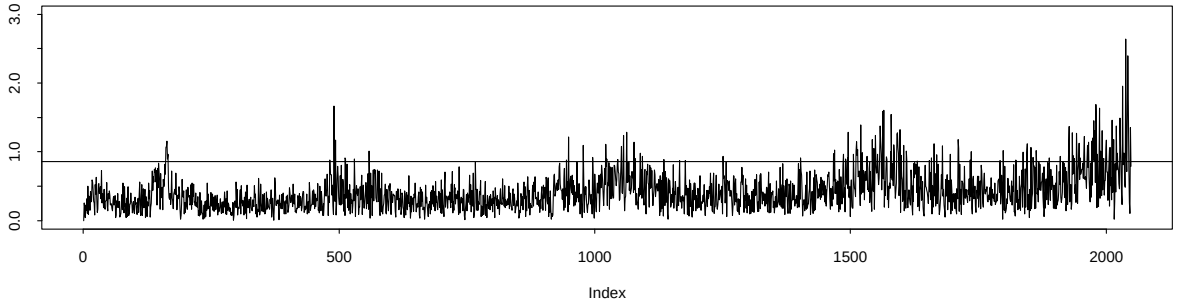

Random Bits Dataset 1 (1-128)
Value
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
120
Index

Power Spectrum: Random Bits Datasets 1-4

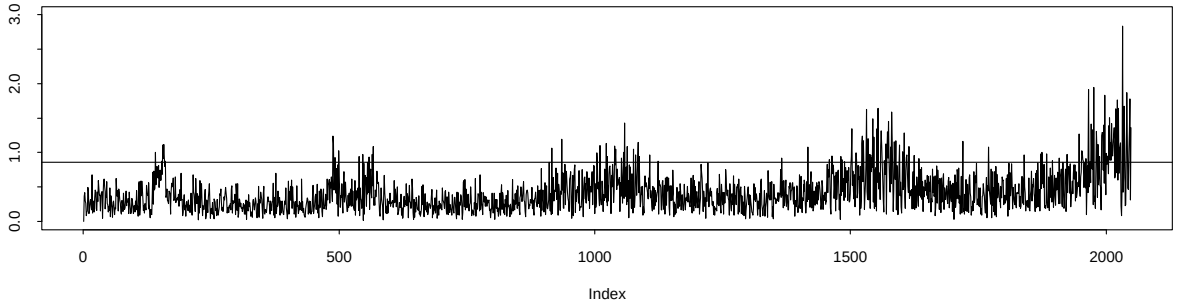
1



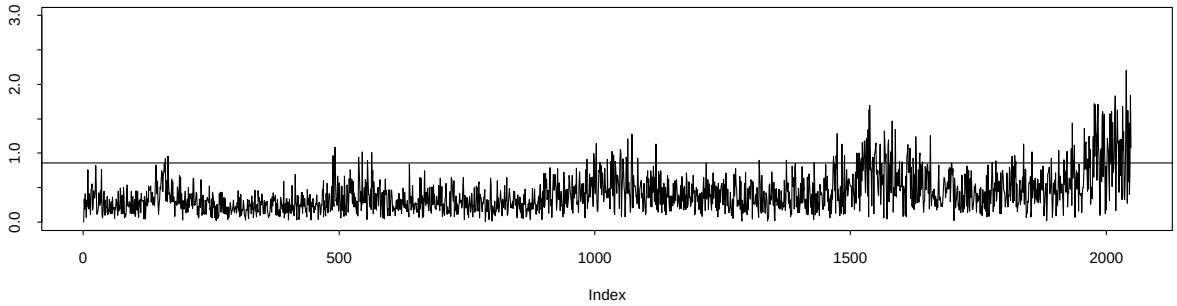
2

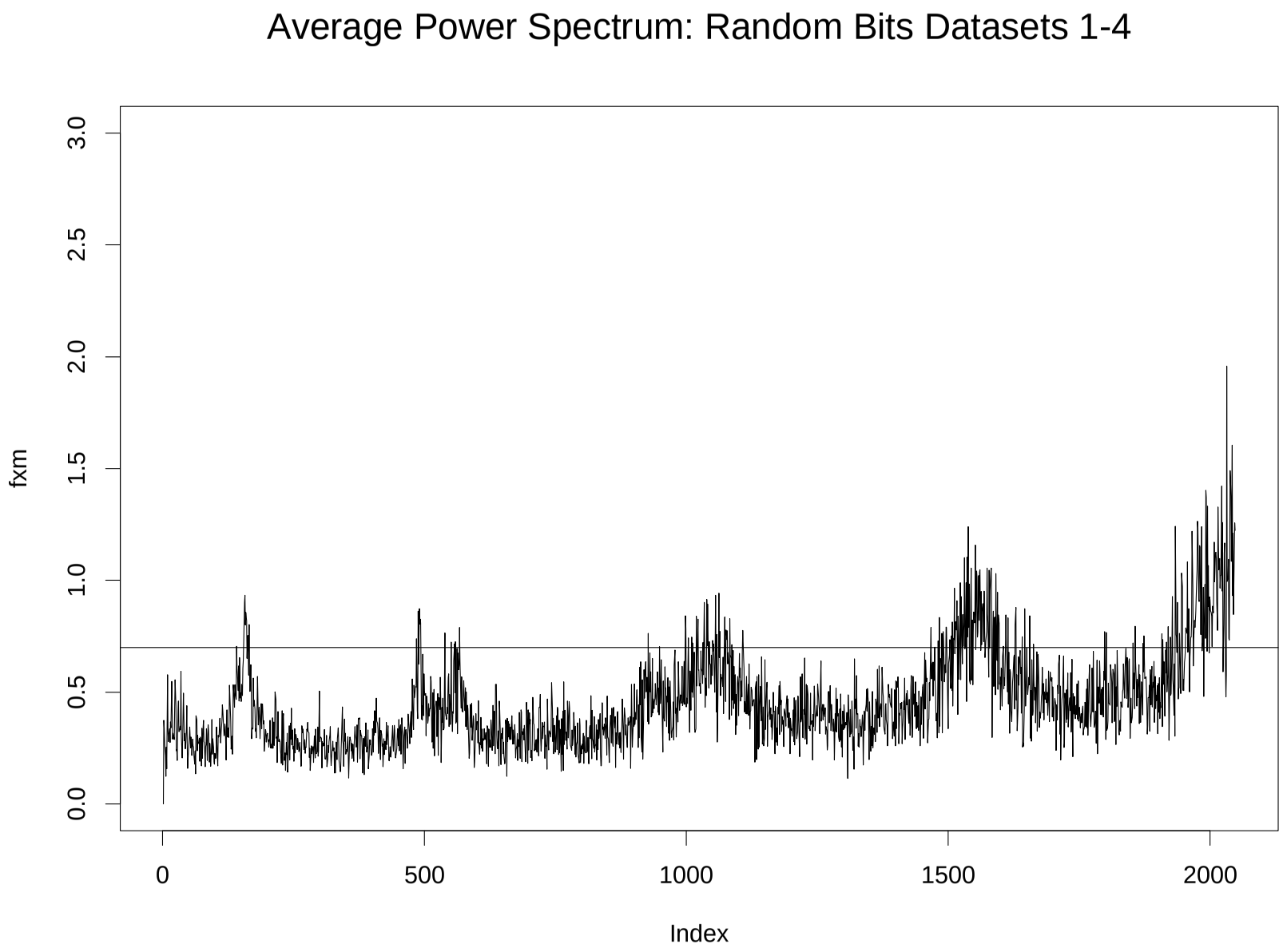


3



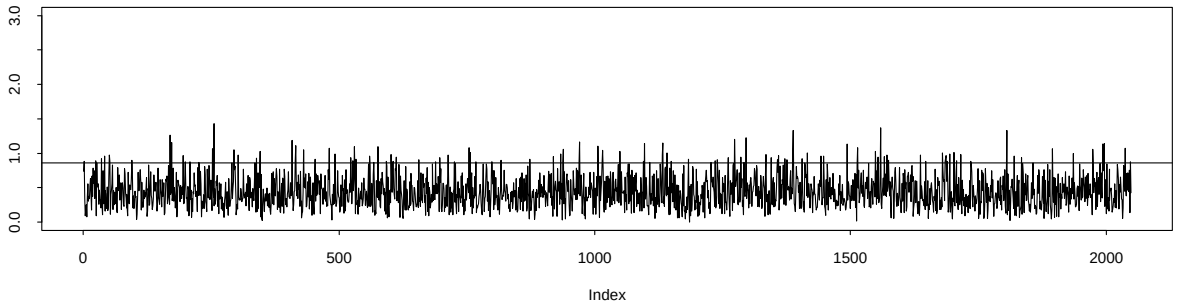
4



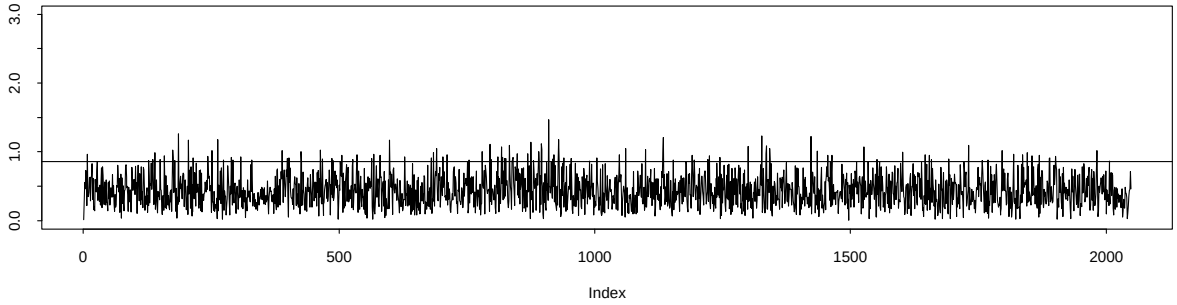

Average Power Spectrum: Random Bits Datasets 1-4
fxm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
500
1000
1500
2000
Index

Power Spectrum: Splus Random Bits

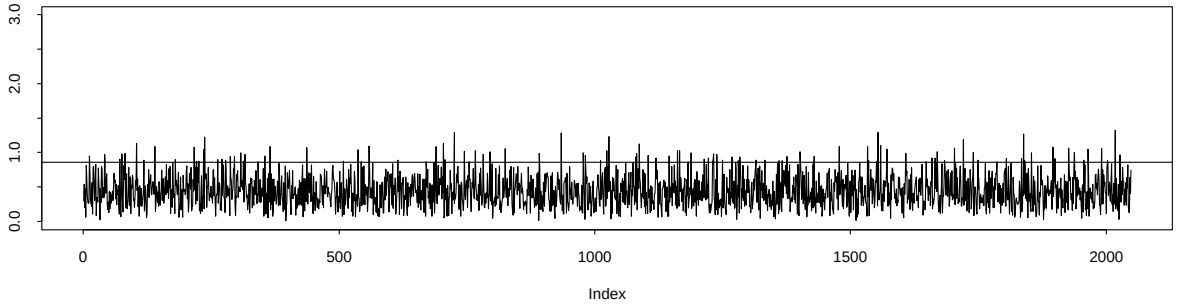
1



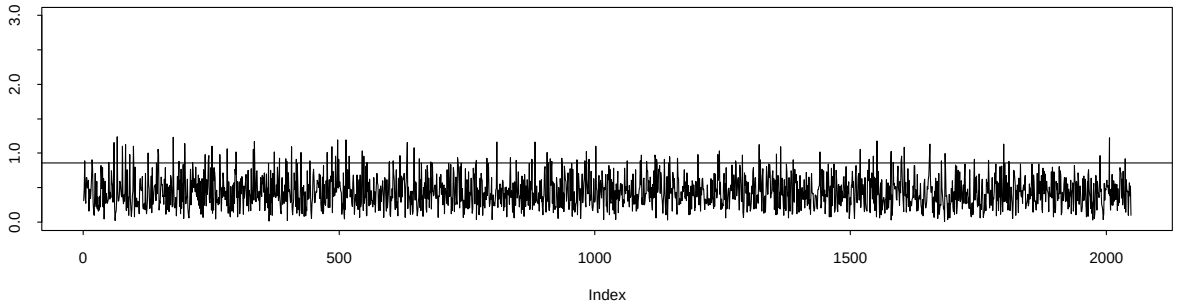
2

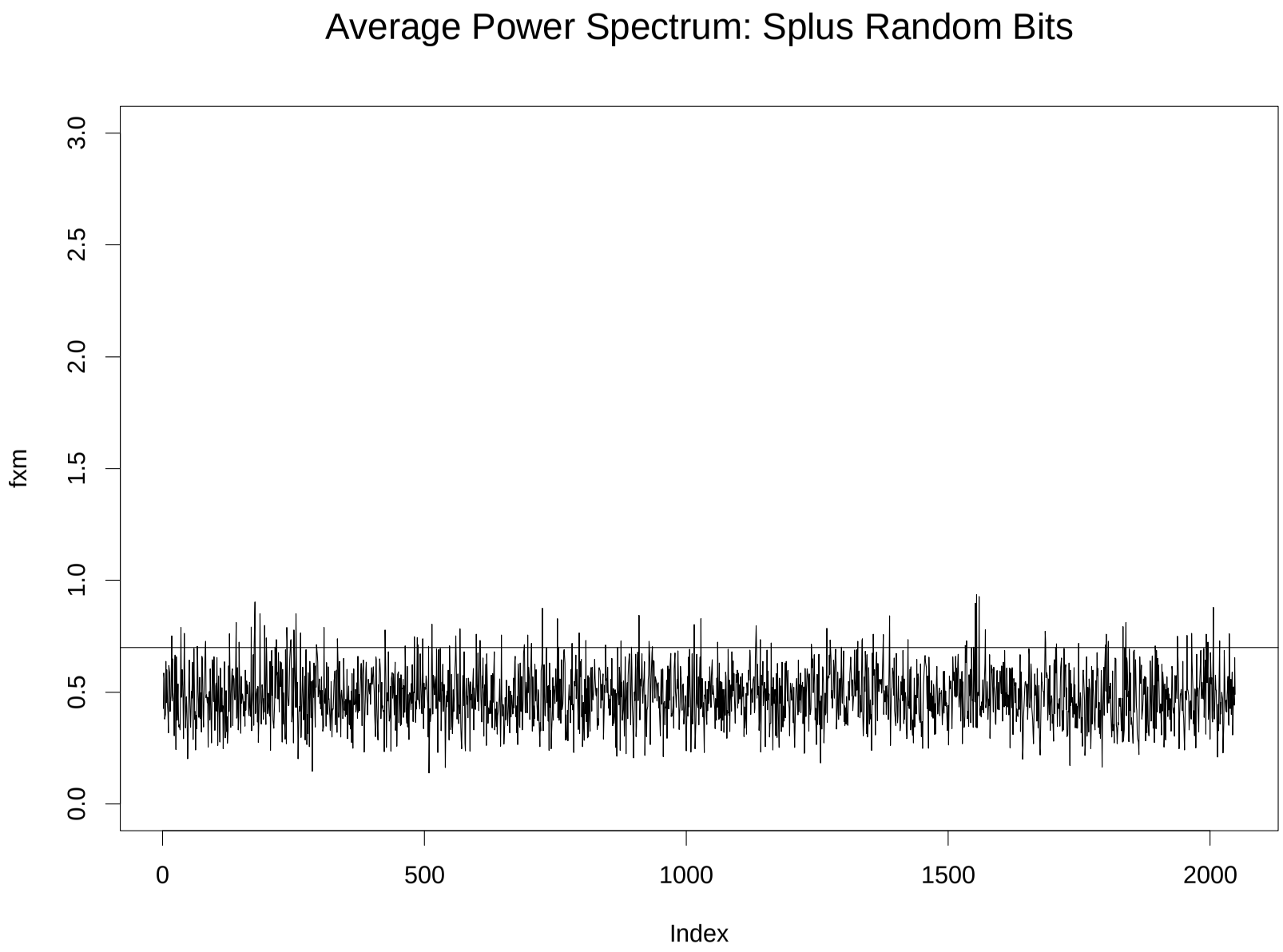


3



4




Average Power Spectrum: Splus Random Bits
fxm
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0
500
1000
1500
2000
Index

2: Mathematical and Statistical Background

a: Vectors and Inner Products

b: Orthogonal Bases

c: Means and Variances of Transformations

2a: Vectors and Inner Products

Vector $\equiv$ An ordered sequence of numbers.
(Your Data!)

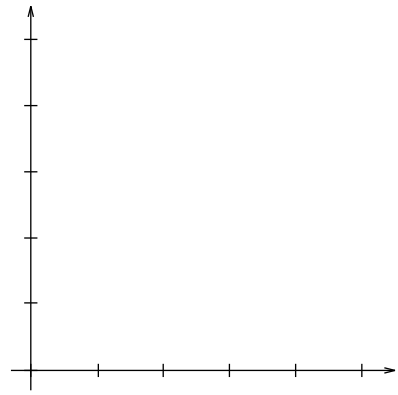
$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \text{Ex: } \boldsymbol{u} = \begin{bmatrix} 5.1 \\ 5.8 \\ 5.9 \\ 5.4 \\ 6.0 \\ 5.3 \end{bmatrix}$$

Notation:
small bold $\equiv$ vector
small unbold $\equiv$ scalar (a number)
large bold $\equiv$ matrix

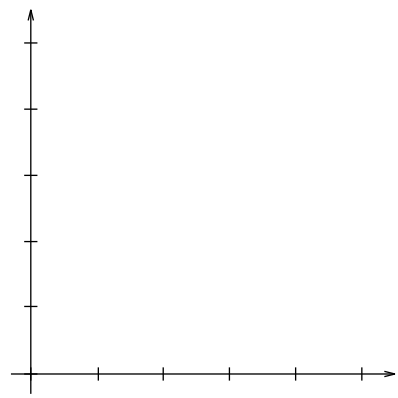
$$\boldsymbol{e} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \qquad \text{length } n$$

"Picturing" Vectors

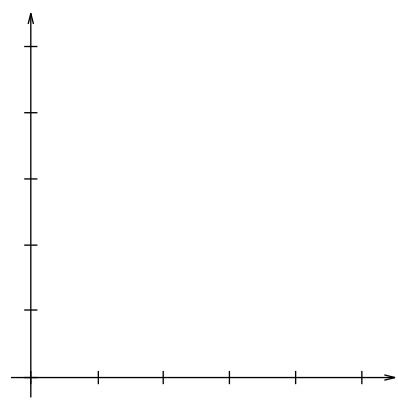
$$\boldsymbol{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$$



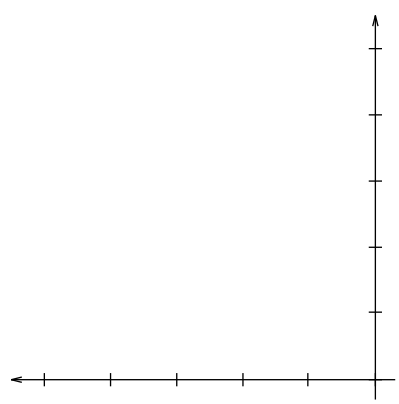
$$\boldsymbol{v} = \begin{bmatrix} 6 \\ 10 \end{bmatrix}$$



$$\boldsymbol{v} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

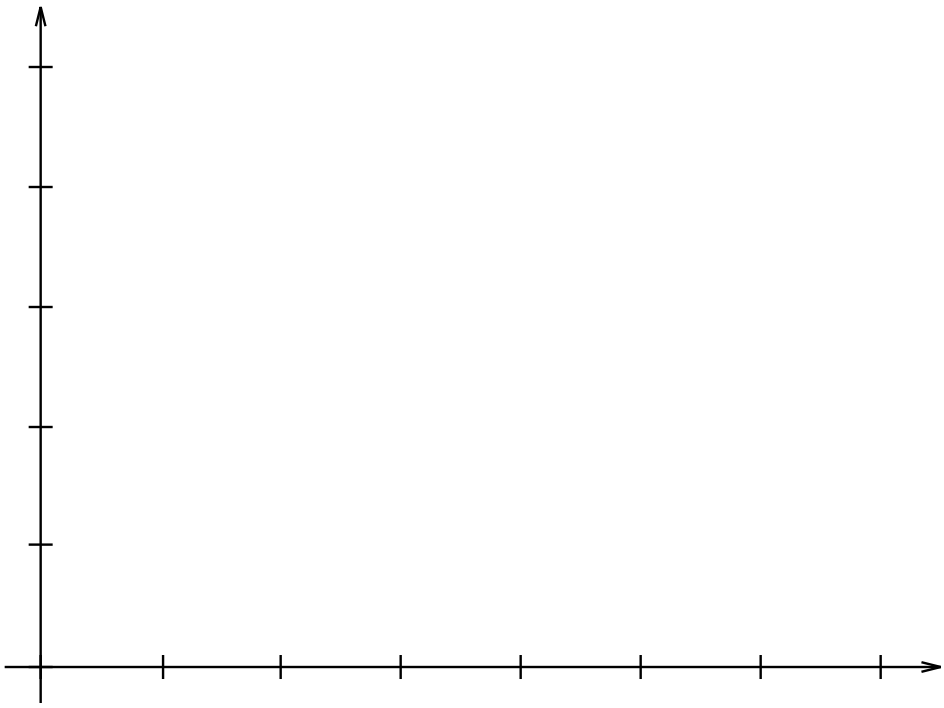


$$\boldsymbol{v} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

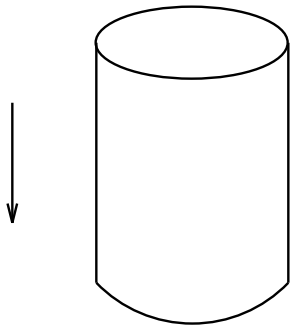


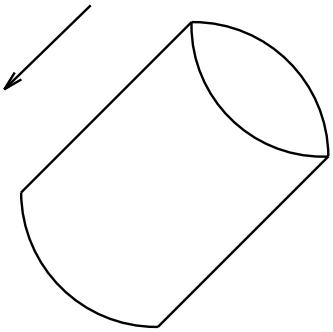
"Picturing" Vectors

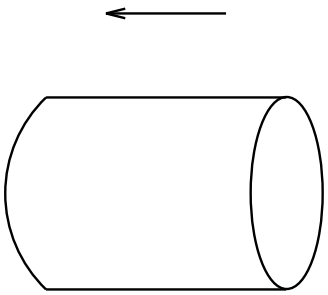
$$\boldsymbol{u} = \begin{bmatrix} 5.1 \\ 5.8 \\ 5.9 \\ 5.4 \\ 6.0 \\ 5.3 \end{bmatrix}$$



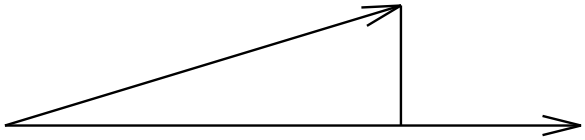
Inner Products

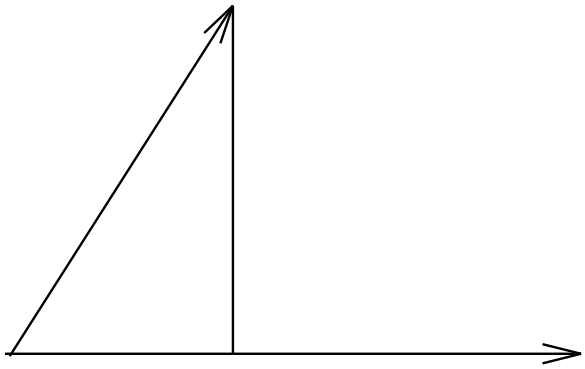






Inner Products





Inner Products

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Length of a Vector $\boldsymbol{u} \equiv \|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$

Inner Product of $\equiv < \boldsymbol{u}, \boldsymbol{v} > = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos\theta$,
where θ is the angle between the two vectors.
(Also called Dot Product and noted with $\boldsymbol{u} \circ \boldsymbol{v}$)

$$< \boldsymbol{u}, \boldsymbol{v} > = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos\theta = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Note: $\|\boldsymbol{v}\| = \sqrt{< \boldsymbol{v}, \boldsymbol{v} >}$

Inner Products

Ex: $\boldsymbol{u} = \begin{bmatrix} 4 \\ 7 \\ -5 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} 6 \\ 3 \\ 3 \\ -4 \end{bmatrix}$

$\|\boldsymbol{u}\| = \sqrt{4^2 + 7^2 + (-5)^2 + 2^2} = 9.70$

$< \boldsymbol{u}, \boldsymbol{v} >= (4)(6) + (7)(3) + (-5)(3) + (2)(-4)$
$= 22$

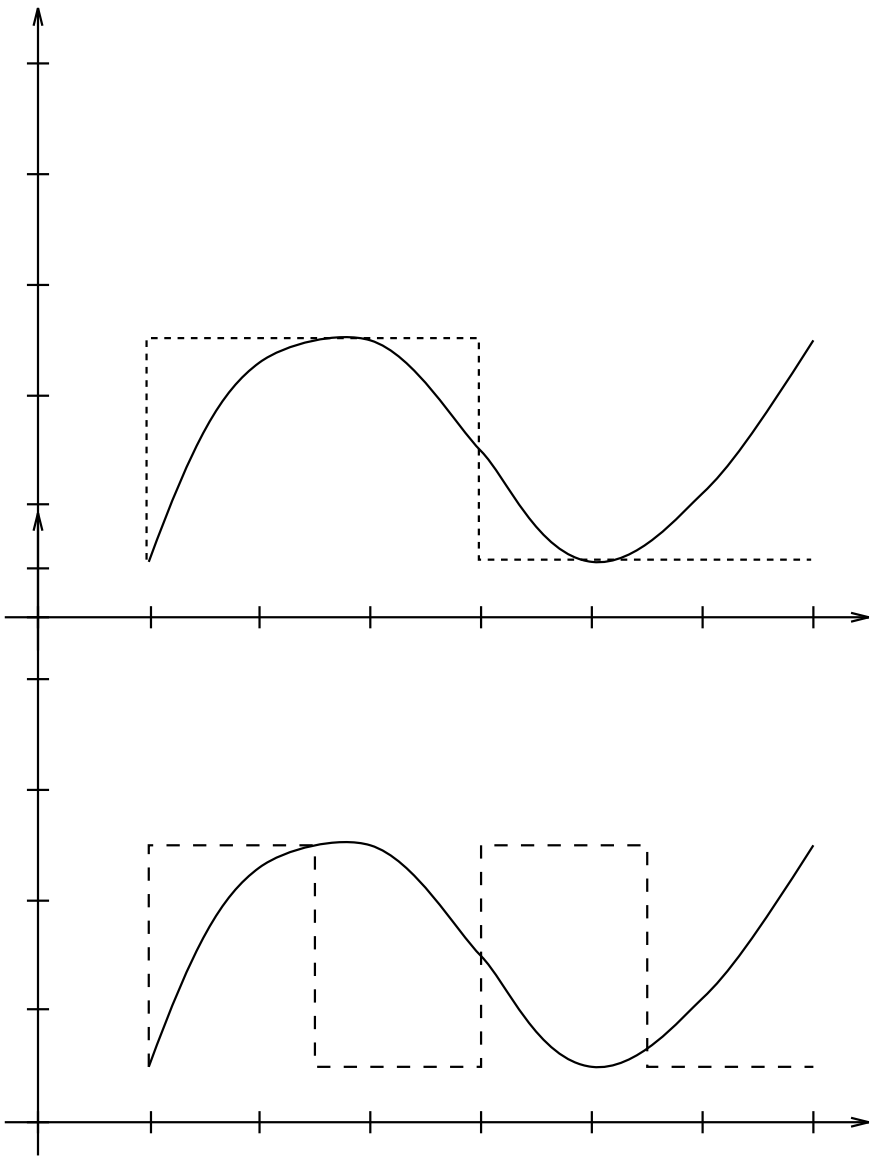
Assume: $\|\boldsymbol{v}\| = 1$,

If $< \boldsymbol{u}, \boldsymbol{v} >$ is large, then $\boldsymbol{u}$ has a large component in the direction of $\boldsymbol{v}$.

If $< \boldsymbol{u}, \boldsymbol{v} >$ is small, then $\boldsymbol{u}$ has a small component in the direction of $\boldsymbol{v}$.

Note: Instead of component, we can say power.

Inner Products



High Power

Low Power

If one vector has the same trend as another it has high power with respect to that vector.

Sample Statistics and Inner Products

Ex: $\boldsymbol{u} = \begin{bmatrix} 4 \\ 7 \\ -5 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} 6 \\ 3 \\ 3 \\ -4 \end{bmatrix}$

Mean$(\boldsymbol{u}) = \overline{u} = \frac{1}{n} \sum u_i$

Variance$(\boldsymbol{u}) = \frac{1}{n} \sum (u_i - \overline{u})^2$

Covariance$(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{n} \sum (u_i - \overline{u})(v_i - \overline{v})$

Note:

Let $\boldsymbol{u} - \overline{u}\boldsymbol{e} = \begin{bmatrix} 4 - \overline{u} \\ 7 - \overline{u} \\ -5 - \overline{u} \\ 2 - \overline{u} \end{bmatrix}$

Variance$(\boldsymbol{u}) = \frac{1}{n} \|\boldsymbol{u} - \overline{u}\boldsymbol{e}\|^2$

Covariance$(\boldsymbol{u}, \boldsymbol{v}) = \frac{1}{n} < \boldsymbol{u} - \overline{u}\boldsymbol{e}, \boldsymbol{v} - \overline{v}\boldsymbol{e} >$

Linear Combinations

Ex: $\boldsymbol{u} = \begin{bmatrix} 4 \\ 7 \\ -5 \\ 2 \end{bmatrix}, \qquad \boldsymbol{v} = \begin{bmatrix} 6 \\ 3 \\ 3 \\ -4 \end{bmatrix}$

$$3\boldsymbol{u} - 2\boldsymbol{v} = \quad 3\begin{bmatrix} 4 \\ 7 \\ -5 \\ 2 \end{bmatrix} - 2\begin{bmatrix} 6 \\ 3 \\ 3 \\ -4 \end{bmatrix} \quad = \quad \begin{bmatrix} 0 \\ 15 \\ -21 \\ ? \end{bmatrix}$$

Bases

Basis in n-dimensional space $\equiv n$ vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ such that any vector $\boldsymbol{v}$ can be represented as a linear combination of the basis vectors.
($\boldsymbol{v} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \cdots + c_n\boldsymbol{b}_n$, for some $c_1, c_2, \cdots, c_n$)

Ex: $n = 3, \quad \boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$$\boldsymbol{u} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix} \quad = \quad 4\boldsymbol{b}_1 - 3\boldsymbol{b}_2 + 2\boldsymbol{b}_3 \quad =$$

$$4\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Orthogonal Bases

Orthogonal Basis in n-dimensional space $\equiv$
a basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ such that

(1) $< \boldsymbol{b}_i, \boldsymbol{b}_j >= 0$ for $i \neq j$ (orthogonal)

(2) $\|\boldsymbol{b}_i\| = 1$ for all i (unit length)

Important Result:
If $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ is an orthogonal basis, then
$\boldsymbol{u} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \cdots + c_n\boldsymbol{b}_n$, where $c_i =< \boldsymbol{u}, \boldsymbol{b}_i >$

Orthogonal Bases

Ex: $\boldsymbol{b}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix}, \boldsymbol{b}_3 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \end{bmatrix}$

$$\boldsymbol{u} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

$$c_1 =< \boldsymbol{u}, \boldsymbol{b}_1 >= \frac{1}{\sqrt{3}}$$
$$c_2 =< \boldsymbol{u}, \boldsymbol{b}_2 >= \frac{2}{\sqrt{2}}$$
$$c_3 =< \boldsymbol{u}, \boldsymbol{b}_3 >= \frac{4}{\sqrt{6}}$$

$$\frac{1}{\sqrt{3}} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} + \frac{2}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{bmatrix} + \frac{4}{\sqrt{6}} \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \end{bmatrix} =$$

Matrix Multiplication

$$\boldsymbol{A} = \begin{bmatrix} 4 & -1 & 3 \\ 2 & 5 & 6 \\ -2 & 4 & -2 \\ 5 & -1 & 0 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} - & \boldsymbol{r}_1 & - \\ - & \boldsymbol{r}_2 & - \\ - & \boldsymbol{r}_3 & - \\ - & \boldsymbol{r}_4 & - \end{bmatrix} \qquad \boldsymbol{r}_1^T = \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}, \text{ etc.}$$

$$\boldsymbol{Ab} =$$

$$\begin{bmatrix} < \boldsymbol{b}, \boldsymbol{r}_1 > \\ < \boldsymbol{b}, \boldsymbol{r}_2 > \\ < \boldsymbol{b}, \boldsymbol{r}_3 > \\ < \boldsymbol{b}, \boldsymbol{r}_4 > \end{bmatrix} =$$

$$\begin{bmatrix} (2)(4) & + & (1)(-1) & + & (-3)(3) \\ (2)(2) & + & (1)(5) & + & (-3)(6) \\ (2)(-2) & + & (1)(4) & + & (-3)(-2) \\ (2)(5) & + & (1)(-1) & + & (-3)(0) \end{bmatrix} = \begin{bmatrix} -2 \\ -9 \\ -6 \\ 9 \end{bmatrix}$$

Matrix Multiplication

$$\boldsymbol{A} = \begin{bmatrix} 4 & -1 & 3 \\ 2 & 5 & 6 \\ -2 & 4 & -2 \\ 5 & -1 & 0 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} | & | & | \\ \boldsymbol{c}_1 & \boldsymbol{c}_2 & \boldsymbol{c}_3 \\ | & | & | \end{bmatrix} \qquad \boldsymbol{c}_1 = \begin{bmatrix} 4 \\ 2 \\ -2 \\ 5 \end{bmatrix}, \text{ etc.}$$

$$\boldsymbol{Ab} = b_1\boldsymbol{a}_1 + b_2\boldsymbol{a}_2 + b_3\boldsymbol{a}_3 =$$

$$2\begin{bmatrix} 4 \\ 2 \\ -2 \\ 5 \end{bmatrix} + 1\begin{bmatrix} -1 \\ 5 \\ 4 \\ -1 \end{bmatrix} - 3\begin{bmatrix} 3 \\ 6 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ -9 \\ -6 \\ 9 \end{bmatrix}$$

Matrix Multiplication

$$\begin{bmatrix} 4 & -1 & 3 \\ 2 & 5 & 6 \\ -2 & 4 & -2 \\ 5 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 0 \\ -3 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 13 \\ -9 & 2 \\ -6 & -6 \\ 9 & 20 \end{bmatrix}$$

$$4 \times 3 \quad 3 \times 2 \quad = \quad 4 \times 2$$

$m \times n \quad n \times p \;=\; m \times p$

$m =$ number of rows

$n =$ number of columns

Orthogonal Transformations

$$\boldsymbol{B} = \begin{bmatrix} - & \boldsymbol{b}_1 & - \\ - & \boldsymbol{b}_2 & - \\ \vdots & \vdots & \vdots \\ - & \boldsymbol{b}_n & - \end{bmatrix}$$

is an orthogonal transformation if $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ is an orthogonal basis in n-dimensions.

By the above result

$$\boldsymbol{B}\boldsymbol{u} = \begin{bmatrix} < \boldsymbol{b}_1, \boldsymbol{u} > \\ < \boldsymbol{b}_2, \boldsymbol{u} > \\ \vdots \\ < \boldsymbol{b}_n, \boldsymbol{u} > \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \boldsymbol{c}$$

$$\boldsymbol{u} = c_1\boldsymbol{b}_1 + c_2\boldsymbol{b}_2 + \cdots + c_n\boldsymbol{b}_n$$

Recall: c_i is large (in magnitude) if $\boldsymbol{u}$ has large power with respect to $\boldsymbol{b}_i$.

The vector $\boldsymbol{c}$ is the spectrum of the signal $\boldsymbol{u}$ with respect to the basis.

Statistical Results

Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ be n repetitions of a p-dimensional vector. Let u_{ij} be the ith element of the jth vector.

$$\text{Ex: } \boldsymbol{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \\ u_{51} \\ u_{61} \end{bmatrix} = \boldsymbol{u} = \begin{bmatrix} 5.1 \\ 5.8 \\ 5.9 \\ 5.4 \\ 6.0 \\ 5.3 \end{bmatrix}, \text{ etc.}$$

$$\boldsymbol{E}(\boldsymbol{u}) = \begin{bmatrix} \overline{u}_{1\cdot} \\ \overline{u}_{2\cdot} \\ \vdots \\ \overline{u}p\cdot \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_j u_{1j} \\ \sum_j u_{2j} \\ \vdots \\ \sum_j u_{pj} \end{bmatrix}$$

$$\sigma_{kl} = \frac{1}{n} \sum_j (u_{kj} - \overline{u}_{k\cdot})(u_{lj} - \overline{u}_{l\cdot})$$

$$\boldsymbol{V}(\boldsymbol{u}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \quad p \times p$$

Statistical Results

$$\boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pn} \end{bmatrix} \quad p \times n$$

[Note: $\boldsymbol{V}(\boldsymbol{u}) = (\boldsymbol{U} - \boldsymbol{E}(\boldsymbol{u})\boldsymbol{e}T)(\boldsymbol{U} - \boldsymbol{E}(\boldsymbol{u})\boldsymbol{e}T)^T$]

Let $\boldsymbol{A}$ be a $p \times p$ matrix.

Let $\boldsymbol{W} = \boldsymbol{A}\boldsymbol{U}$.

Then

$$\boldsymbol{E}(\boldsymbol{w}) = \boldsymbol{A}\boldsymbol{E}(\boldsymbol{u})$$

$$\boldsymbol{V}(\boldsymbol{w}) = \boldsymbol{A}\boldsymbol{V}(\boldsymbol{u})\boldsymbol{A}^T$$

White Noise

$\boldsymbol{u}$ is white noise if

$$\boldsymbol{E}(\boldsymbol{u}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and } \boldsymbol{V}(\boldsymbol{u}) = \sigma^2 \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & & \\ & & & 1 \end{bmatrix}$$

Important Result:

If $\boldsymbol{u}$ is white noise and $\boldsymbol{A}$ is a $p \times p$ orthogonal matrix, then $\boldsymbol{w}$, where $\boldsymbol{W} = \boldsymbol{A}\boldsymbol{U}$ is white noise.

$\boldsymbol{w}$ is the spectrum with respect to the basis of the rows of $\boldsymbol{A}$

3: Fourier Analysis

a: Definition

b: Properties

c: Random Bits Example

Definition

Let the sequence $u_0, u_1, \cdots, u_{n-1}$ be a signal observed at n points.

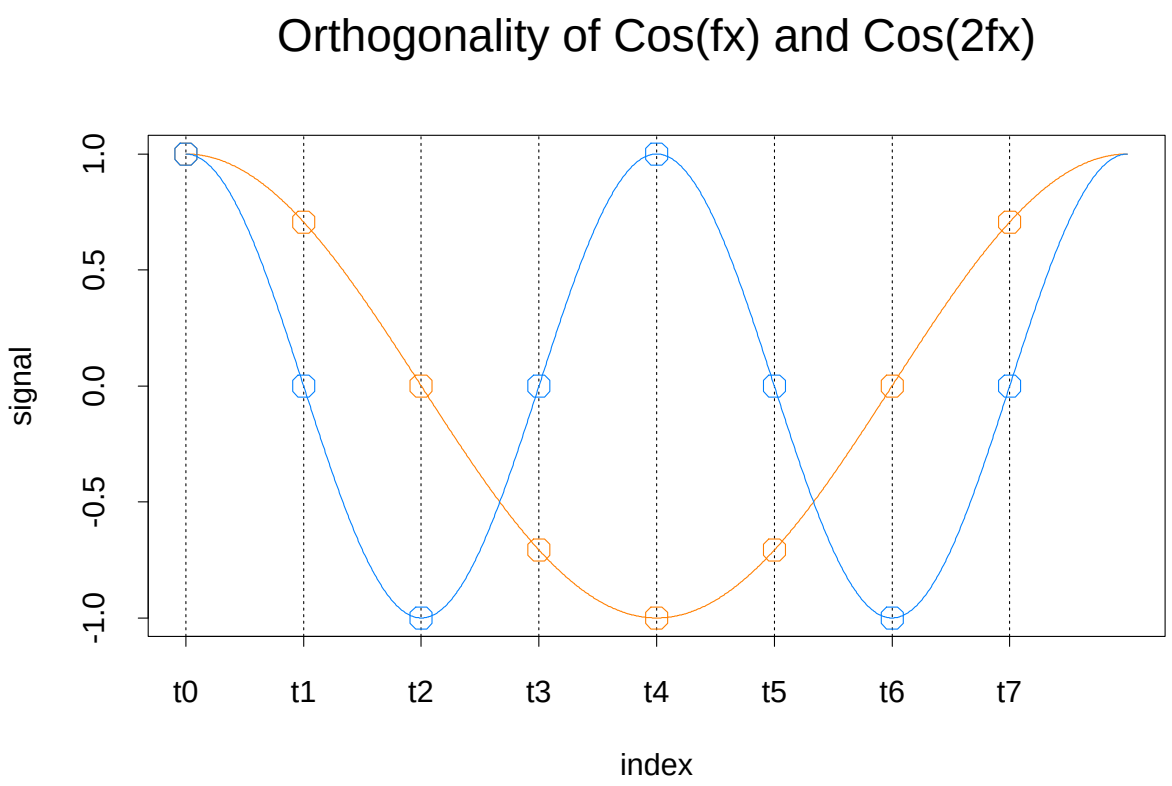
Exs:
(1) u_i is a voltage observed at time i.
(2) u_i is the height of a surface at distance i.

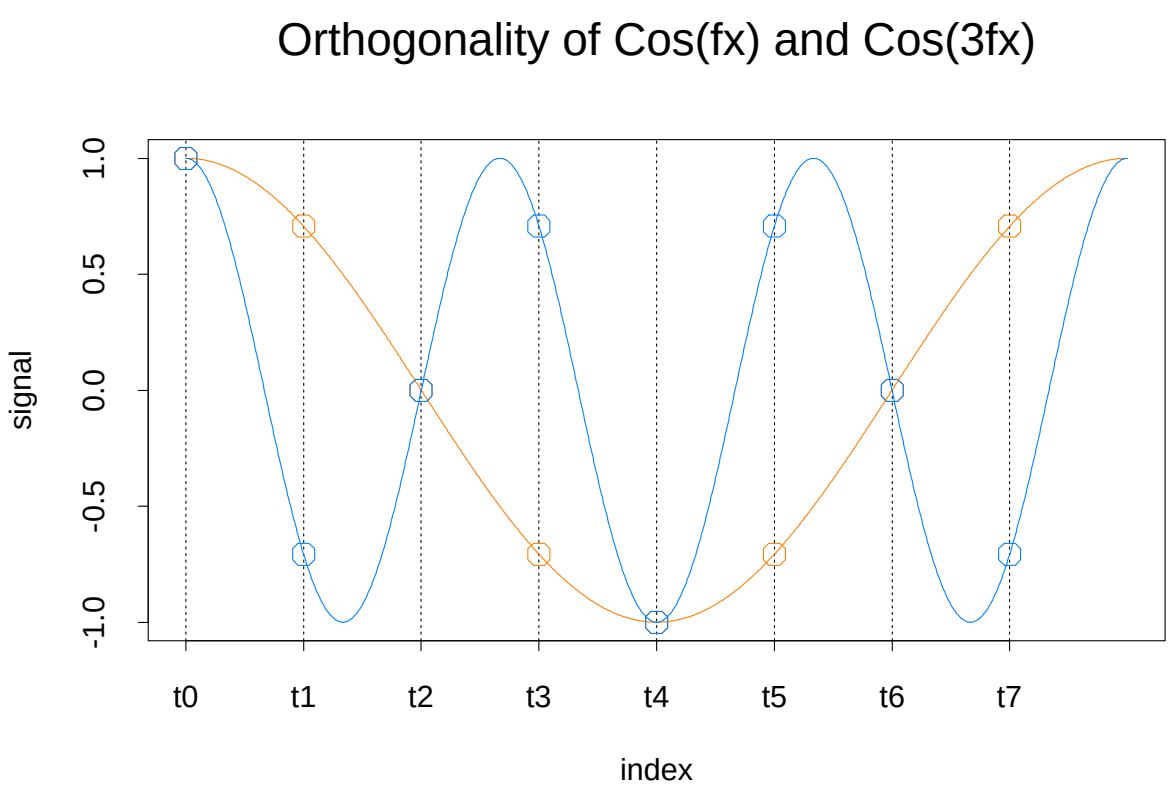
The *Discrete Fourier Transform* (DFT) is a sequence $f_0, f_1, \cdots, f_{n-1}$ where

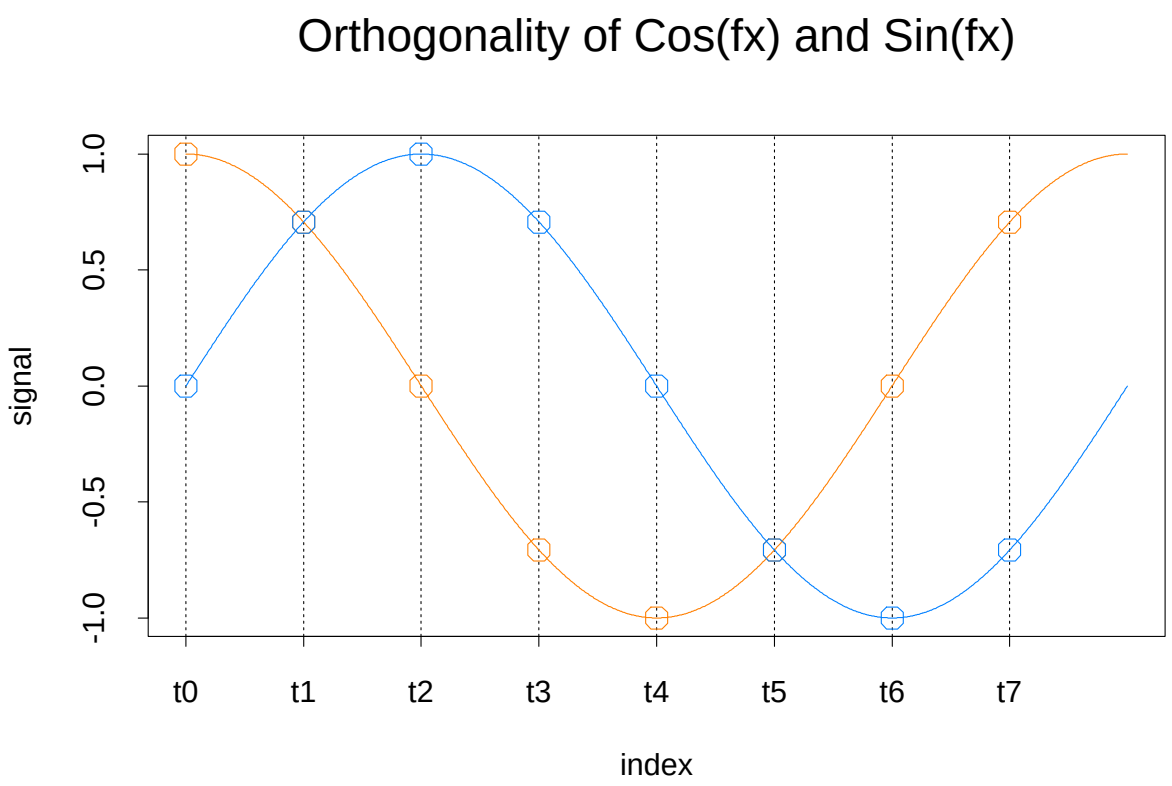
$$f_k = \sum_{j=0}^{n-1} u_j \exp\left(-2\pi i k j/n\right)$$

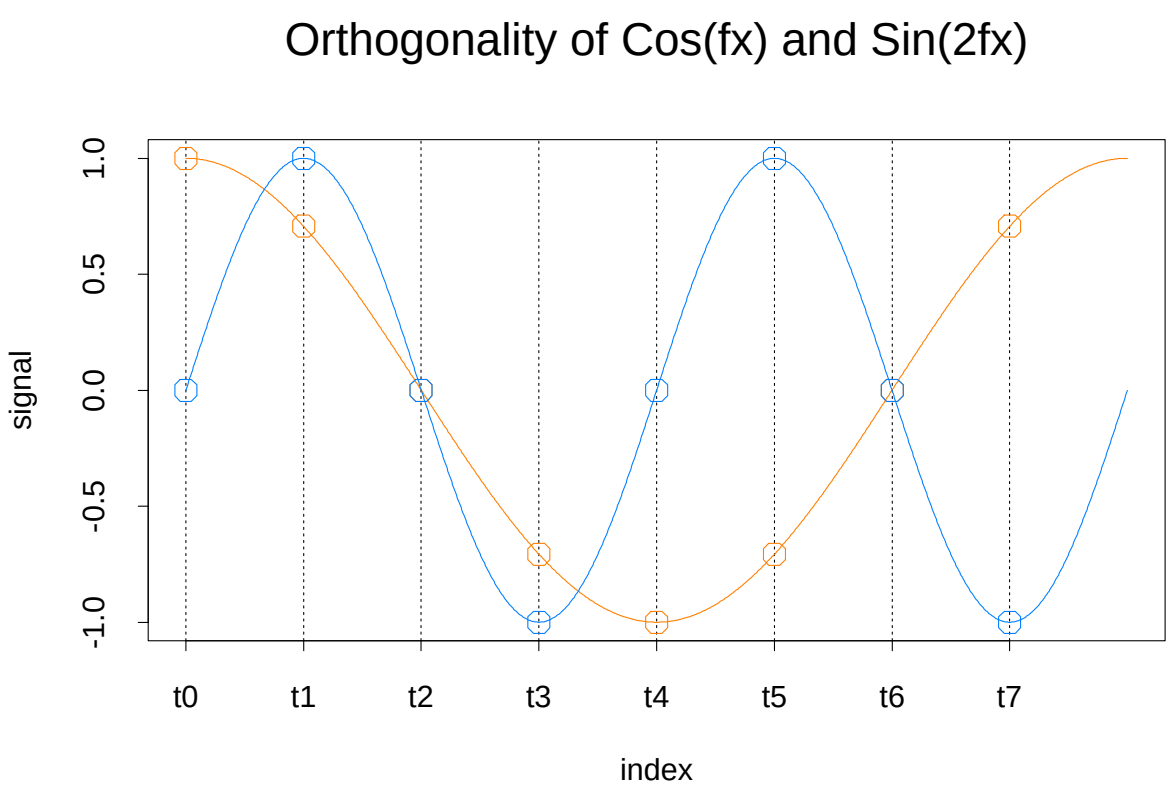
$\exp(ix) = \cos(x) + i \sin(x)$ and $i = \sqrt{-1}$.

Exs:
(1) Units of f_k are cycles per second
(2) Units of f_k are cycles per length


Orthogonality of Cos(fx) and Cos(2fx)
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
t6
t7
index


Orthogonality of Cos(fx) and Cos(3fx)
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
t6
t7
index


Orthogonality of Cos(fx) and Sin(fx)
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
t6
t7
index


Orthogonality of Cos(fx) and Sin(2fx)
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
t6
t7
index

Properties

$$f_k = \sum_{j=0}^{n-1} u_j \exp\left(-2\pi i k j/n\right)$$

Let $\boldsymbol{u} = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}$, $\boldsymbol{f} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_{n-1} \end{bmatrix}$

$$\boldsymbol{T} = \begin{bmatrix} - & \boldsymbol{t}_0 & - \\ - & \boldsymbol{t}_2 & - \\ \vdots & \vdots & \vdots \\ - & \boldsymbol{t}_{n-1} & - \end{bmatrix}$$

$$\boldsymbol{t}_k = \begin{bmatrix} \exp\left(-2\pi i k(0)/n\right) \\ \exp\left(-2\pi i k(1)/n\right) \\ \vdots \\ \exp\left(-2\pi i k(n-1)/n\right) \end{bmatrix}, \text{ etc.}$$

$$\boldsymbol{f} = \boldsymbol{T}\boldsymbol{u}$$

The DFT is an orthogonal transformation, where $\boldsymbol{f}$ is the spectrum with respect to the basis.

Properties

Why do we need sines and cosines?

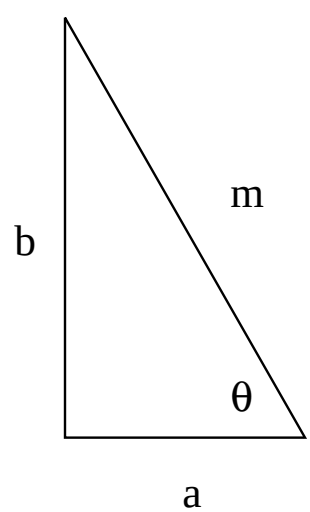
$$f_k = a_k + ib_k$$

$$a_k = \sum_{j=0}^{n-1} u_j \cos(-2\pi kj/n)$$

$$b_k = \sum_{j=0}^{n-1} u_j \sin(-2\pi kj/n)$$

Let $m_k = \sqrt{a_k^2 + b_k^2}$.

$a_k = m_k \cos(\theta)$
$b_k = m_k \sin(\theta)$
$\tan(\theta_k) = b_k/a_k$



Properties

$a_k \cos(-2\pi kj/n) + b_k \sin(-2\pi kj/n) =$

$m_k[\cos(\theta_k)\cos(-2\pi kj/n) + \sin(\theta_k)\sin(-2\pi kj/n) =$

$m_k \cos(-2\pi kj/n - \theta_k)$

m_k is the magnitude of the frequency.
θ_k is the phase of the frequency.

Note: n values $u_0, u_1, \cdots, u_{n-1} \Longrightarrow$
n complex valued $f_0, f_1, \cdots, f_{n-1}$.
(2 n real values)

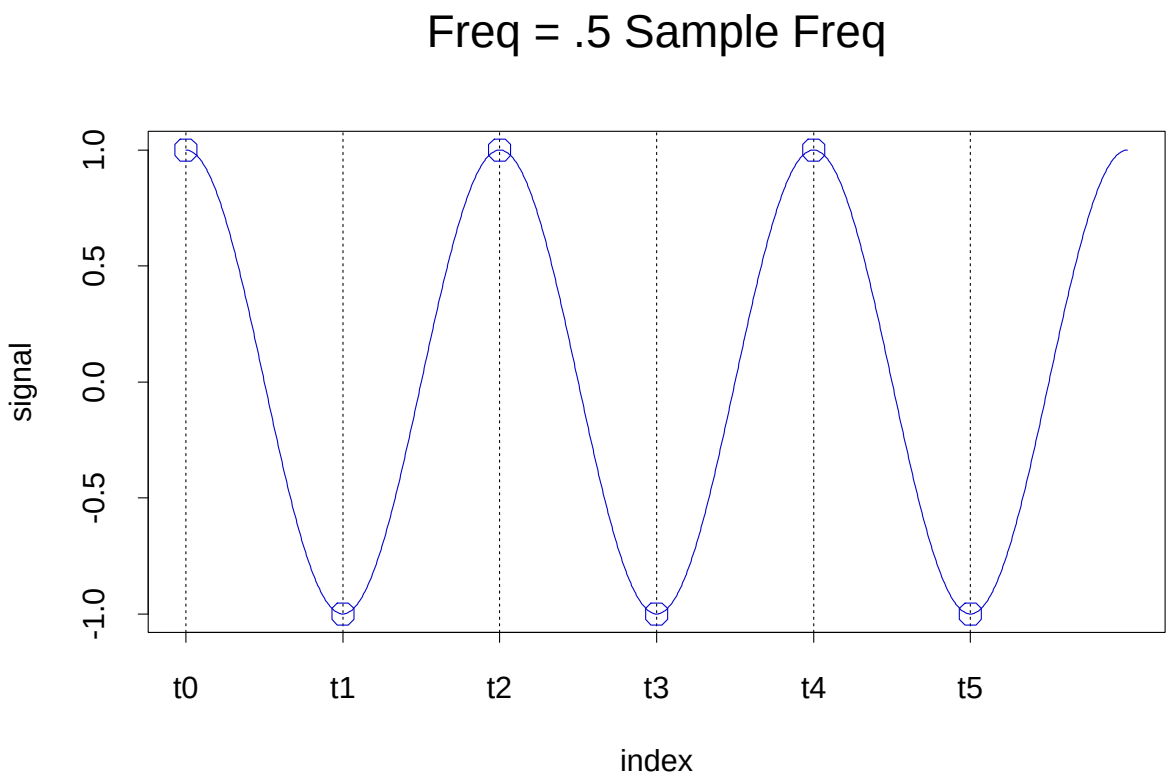
Aliasing

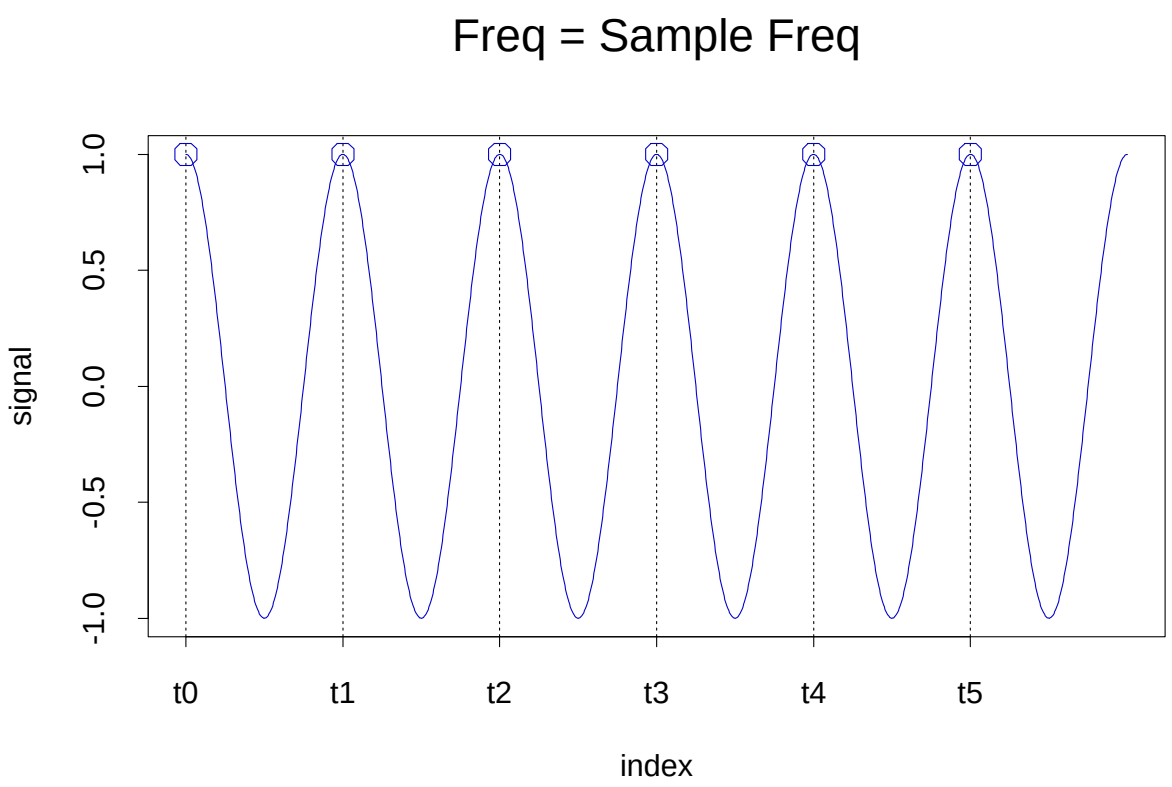
Let D be the time (or distance) from the observation of u_0 and u_n.

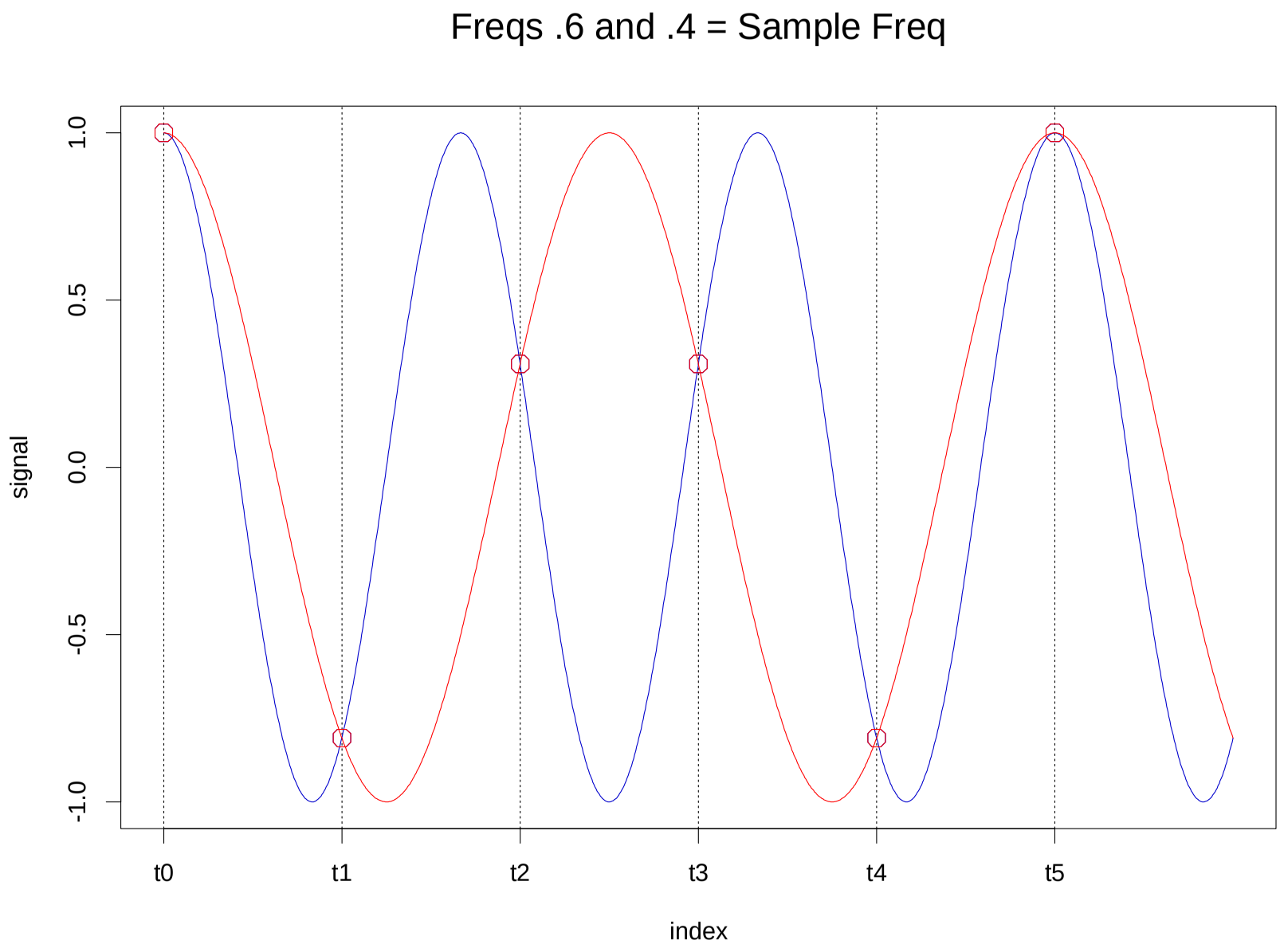
f_k corresponds to k cycles per D.

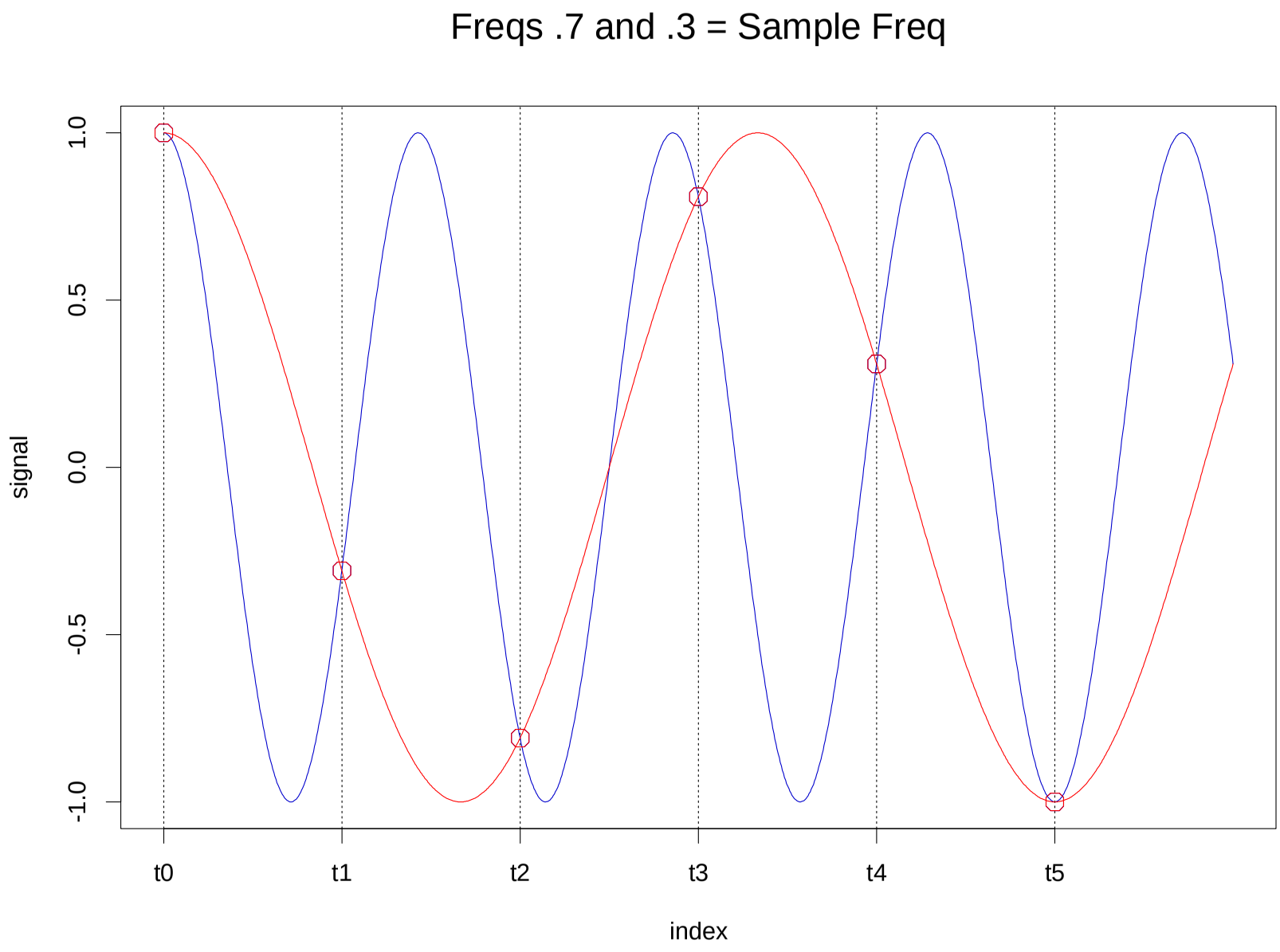
The largest k that can be estimated is $n/2$.
$\equiv$ two observations per period of the wave.
It is called the *Nyquist Frequency*, η

A component with $k > \eta$ appears as a contribution to the component k mod η.


Freq = .5 Sample Freq
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
index


Freq = Sample Freq
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
index


Freqs .6 and .4 = Sample Freq
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
index


Freqs .7 and .3 = Sample Freq
signal
1.0
0.5
0.0
-0.5
-1.0
t0
t1
t2
t3
t4
t5
index

4: Wavelet Analysis

a: Definition

b: Properties

c: Turbulence Study Example

Definition

Let the sequence $u_0, u_1, \cdots, u_{n-1}$ be a signal observed at n points.

Exs:
(1) u_i is a voltage observed at time i.
(2) u_i is the height of a surface at distance i.

The *Discrete Wavelet Transform* (DWT) is a double index sequence d_{jk}, where

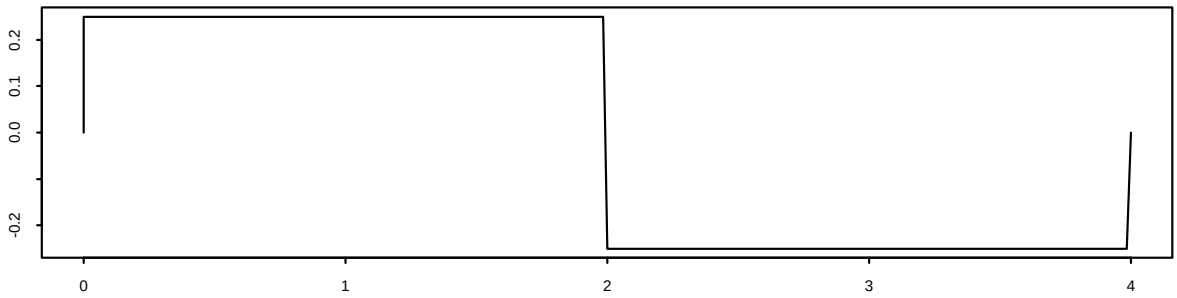
$$d_{jk} = \sum_{i=0}^{n-1} u_i w(2^j(u_i - k))$$

The function $w()$ is called the *mother wavelet*.

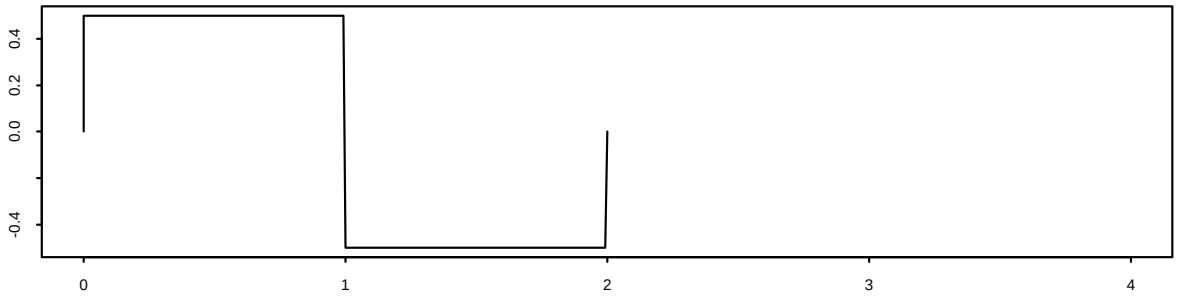
j scales the mother wavelet.
k translates the mother wavelet.

The functions $w(2^j(u_i - k))$ are orthogonal for distinct values of j or k.

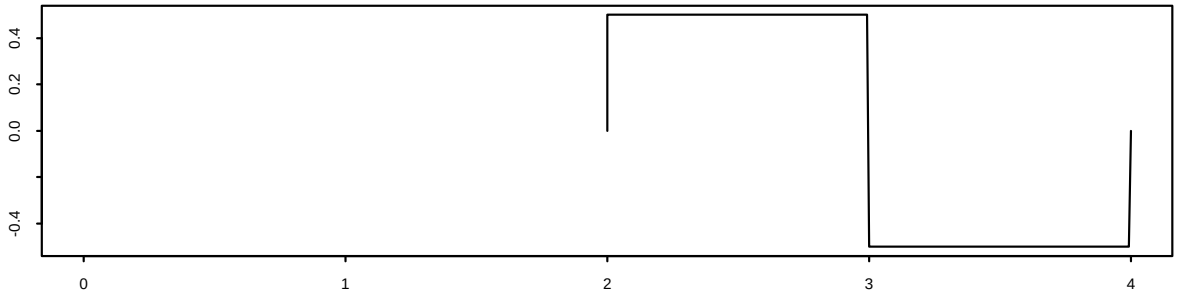
Haar Level 2, Shift 0



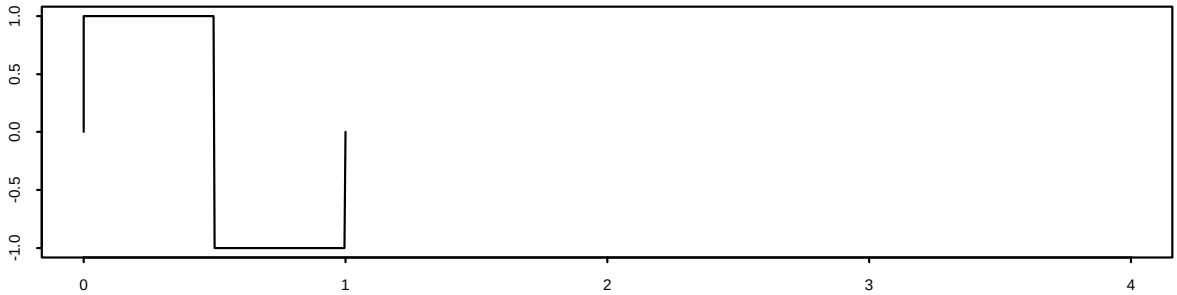
Haar Level 1, Shift 0



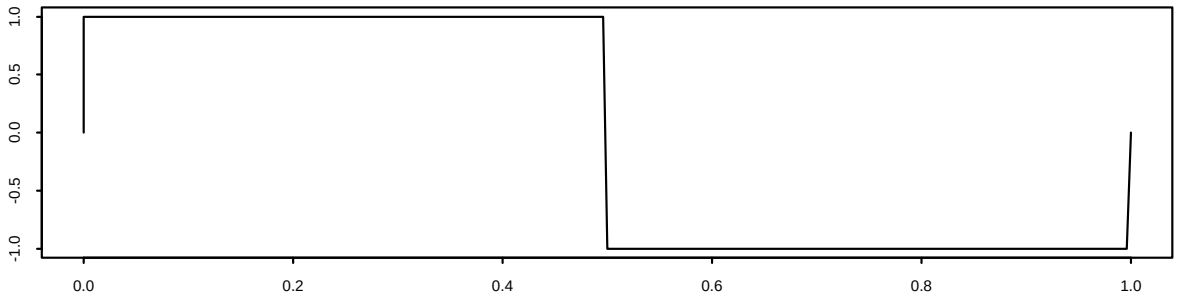
Haar Level 1, Shift 1



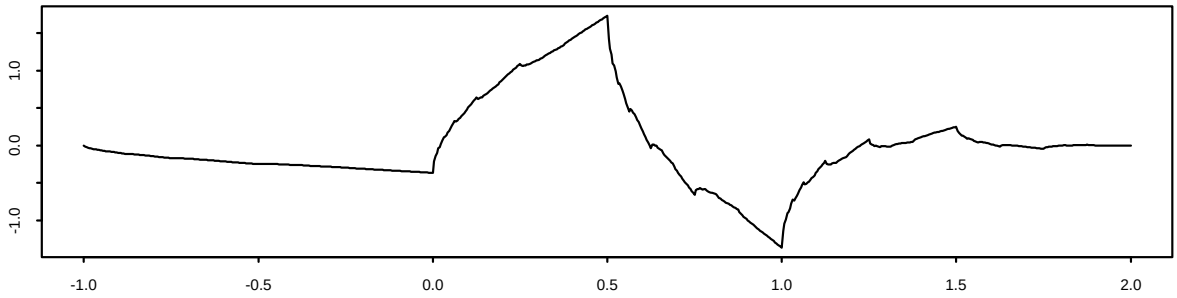
Haar Level 0, Shift 0



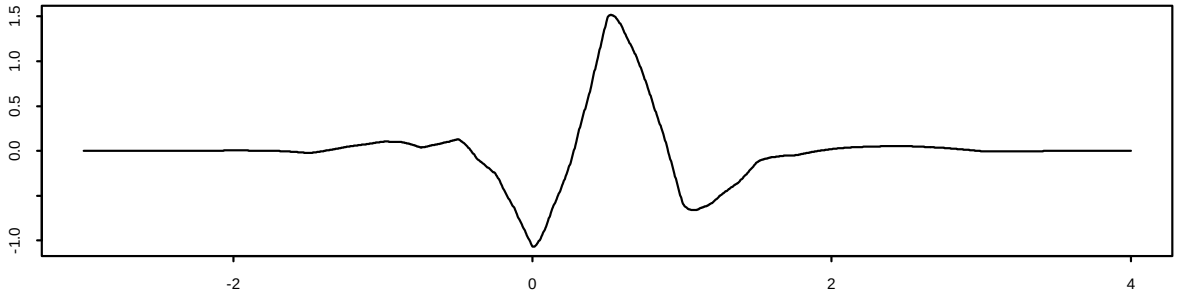
Haar: Mother Wavelet



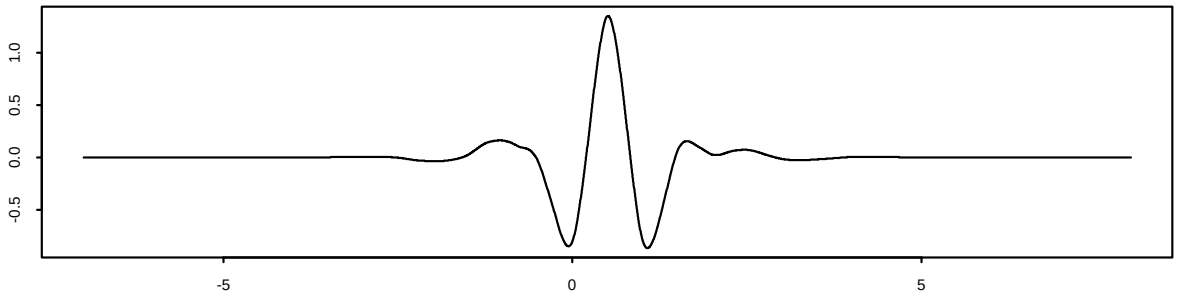
Daublet D4: Mother Wavelet



Symmlet S8: Mother Wavelet



Symmlet S16: Mother Wavelet



More Definition

The *Discrete Wavelet Transform* (DWT) is two double index sequences: one fine scale d_{jk} and one coarse scale s_{jk}, where

$$d_{jk} = \sum_{i=0}^{n-1} u_i w_m(2^j(u_i - k))$$

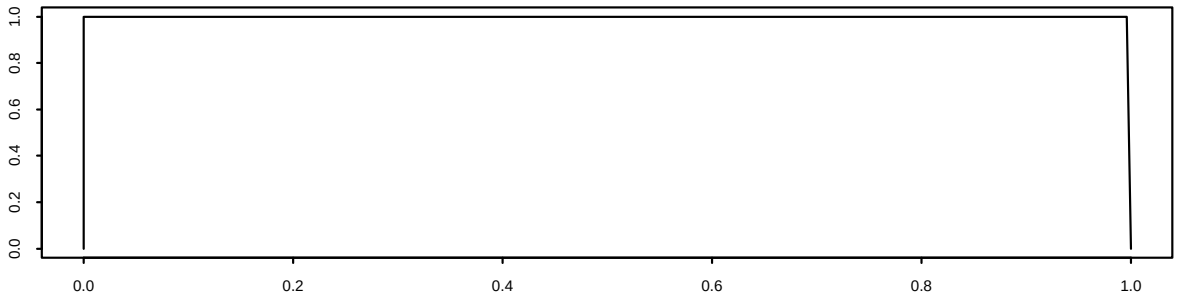
$$s_{jk} = \sum_{i=0}^{n-1} u_i w_f(2^j(u_i - k))$$

The function $w_m()$ is called the *mother wavelet*. The function $w_f()$ is called the *father wavelet*.

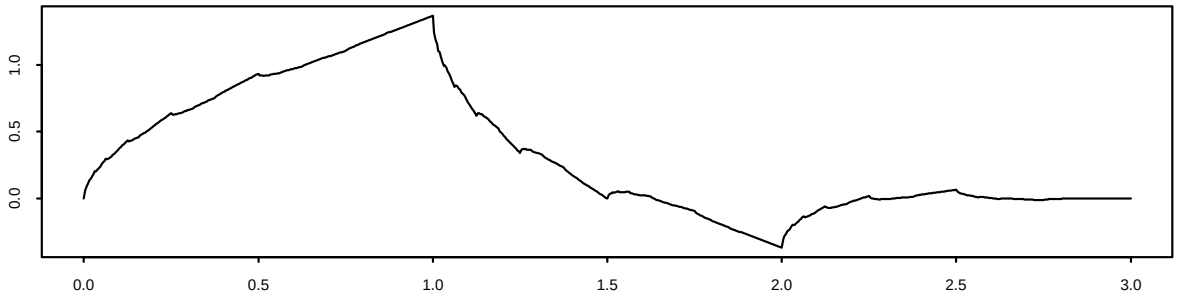
d_{jk} is sensitive to fine scale.
s_{jk} is sensitive to coarse scale.

The functions $w_m(2^j(u_i - k))$ and $w_f(2^j(u_i - k))$ are orthogonal.

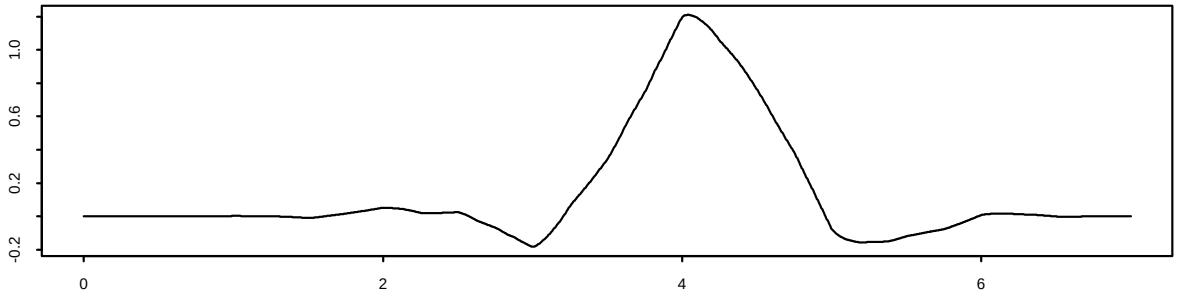
Haar: Father Wavelet



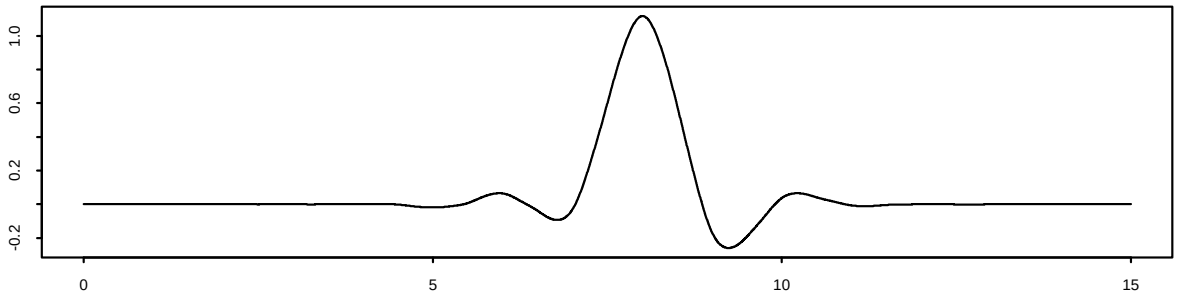
Daublet D4: Father Wavelet

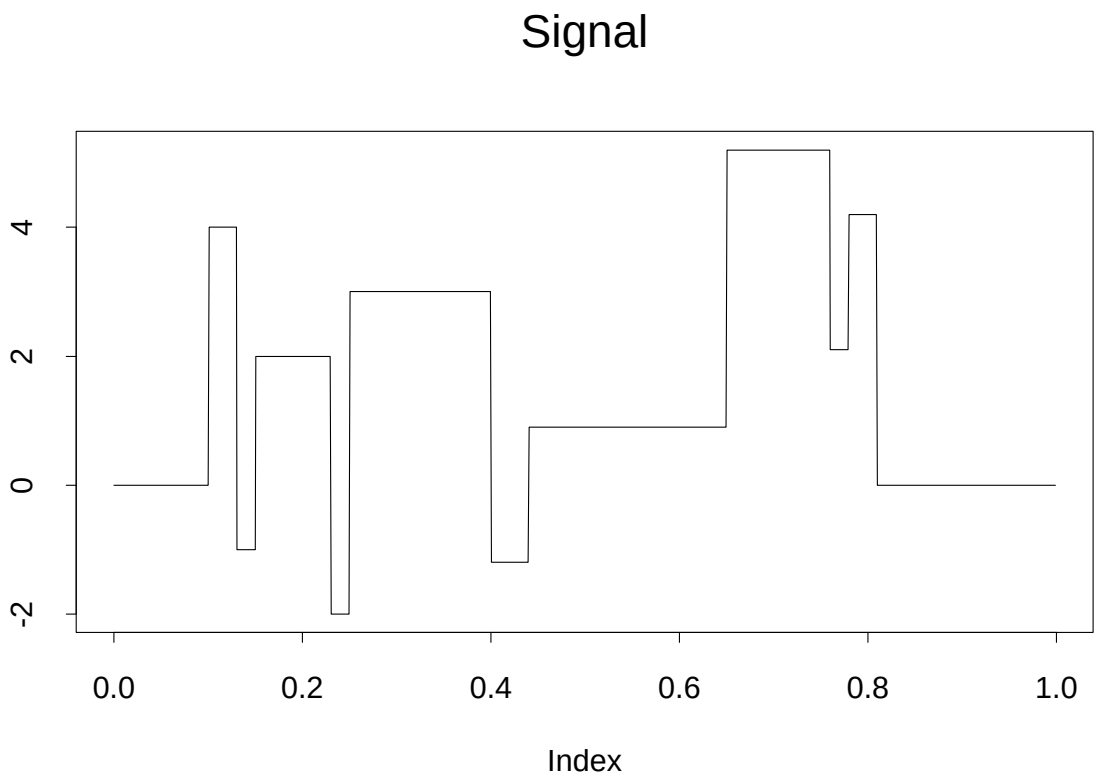


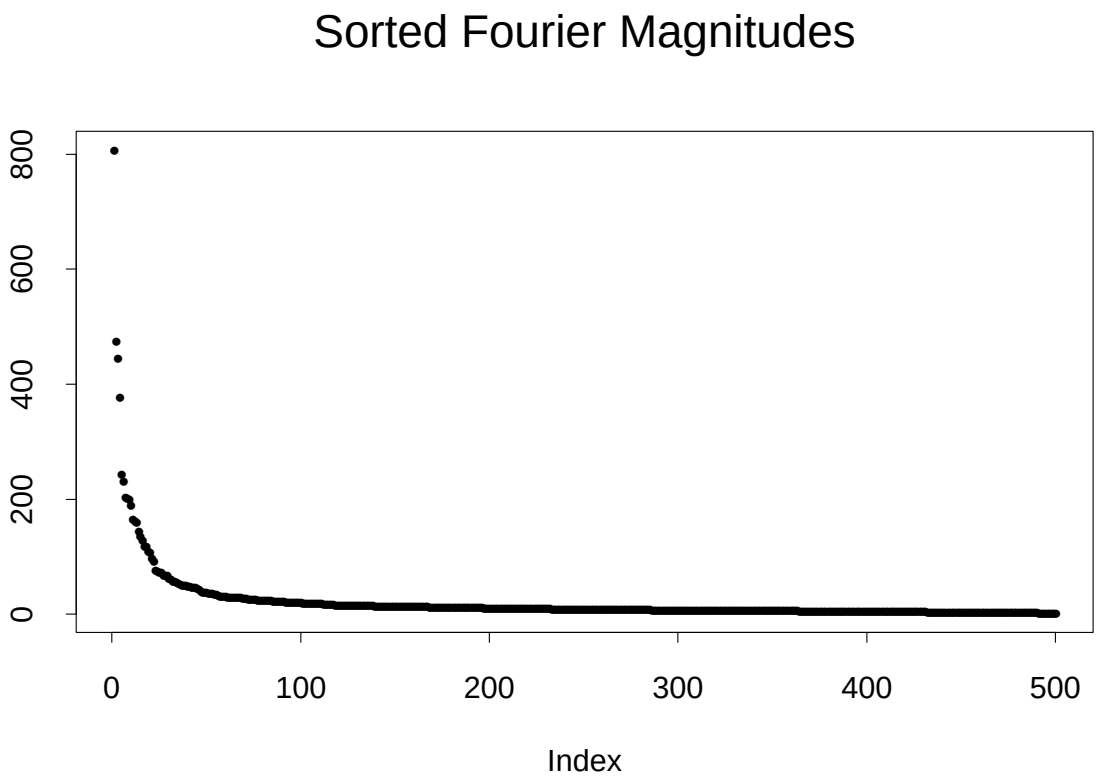
Symmlet S8: Father Wavelet

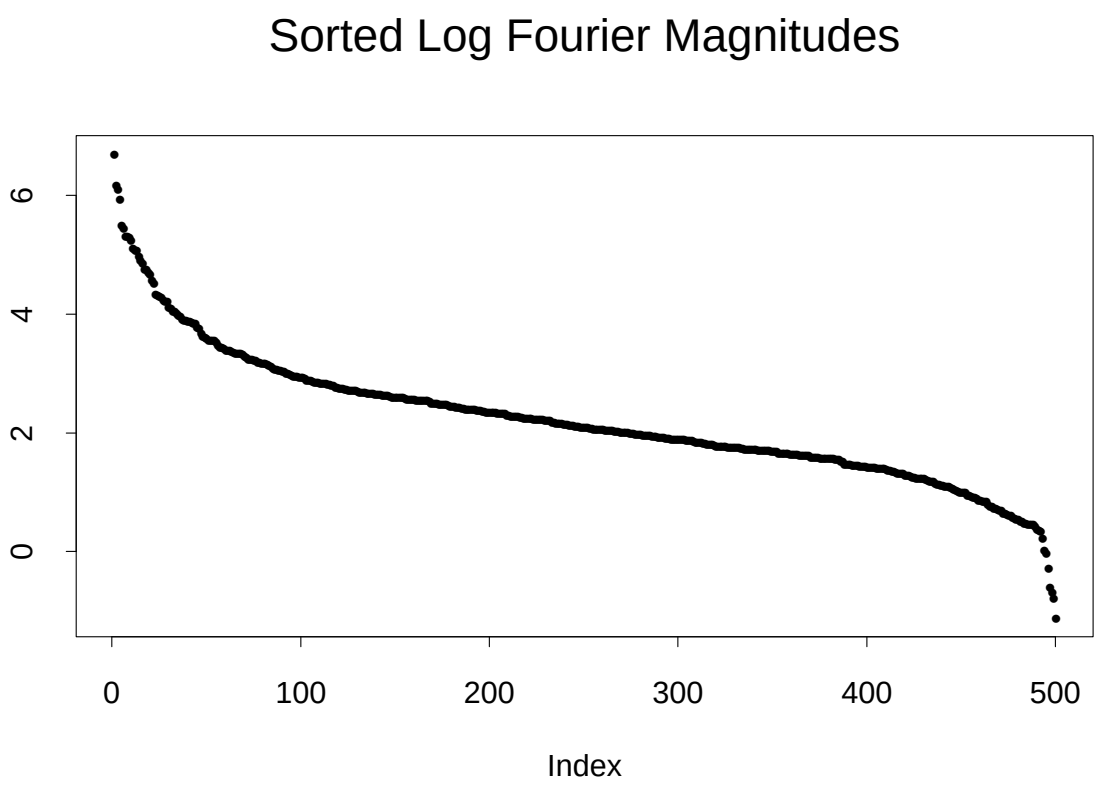


Symmlet S16: Father Wavelet

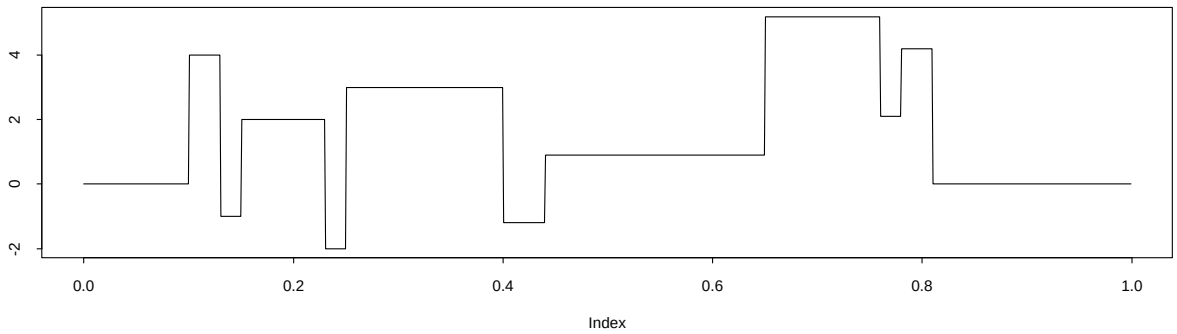



Signal
4
2
0
-2
0.0
0.2
0.4
0.6
0.8
1.0
Index

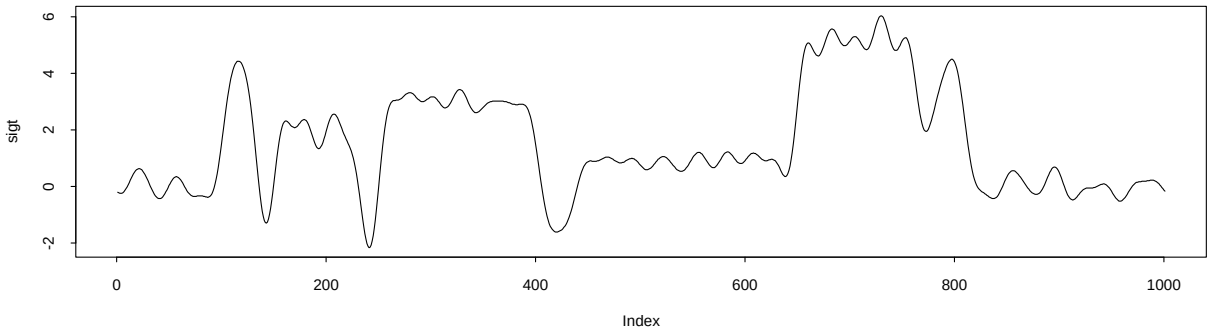

Sorted Fourier Magnitudes
0
200
400
600
800
0
100
200
300
400
500
Index


Sorted Log Fourier Magnitudes
0
2
4
6
0
100
200
300
400
500
Index

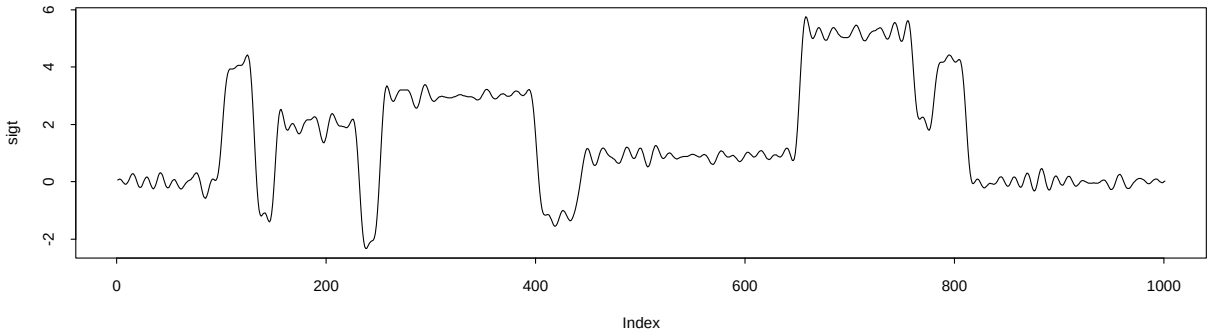
Signal



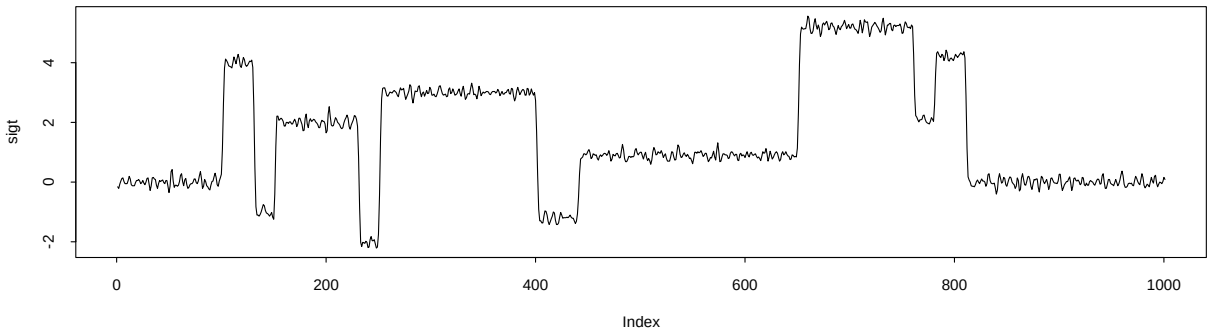
Fourier Signal Reconstruction 5%

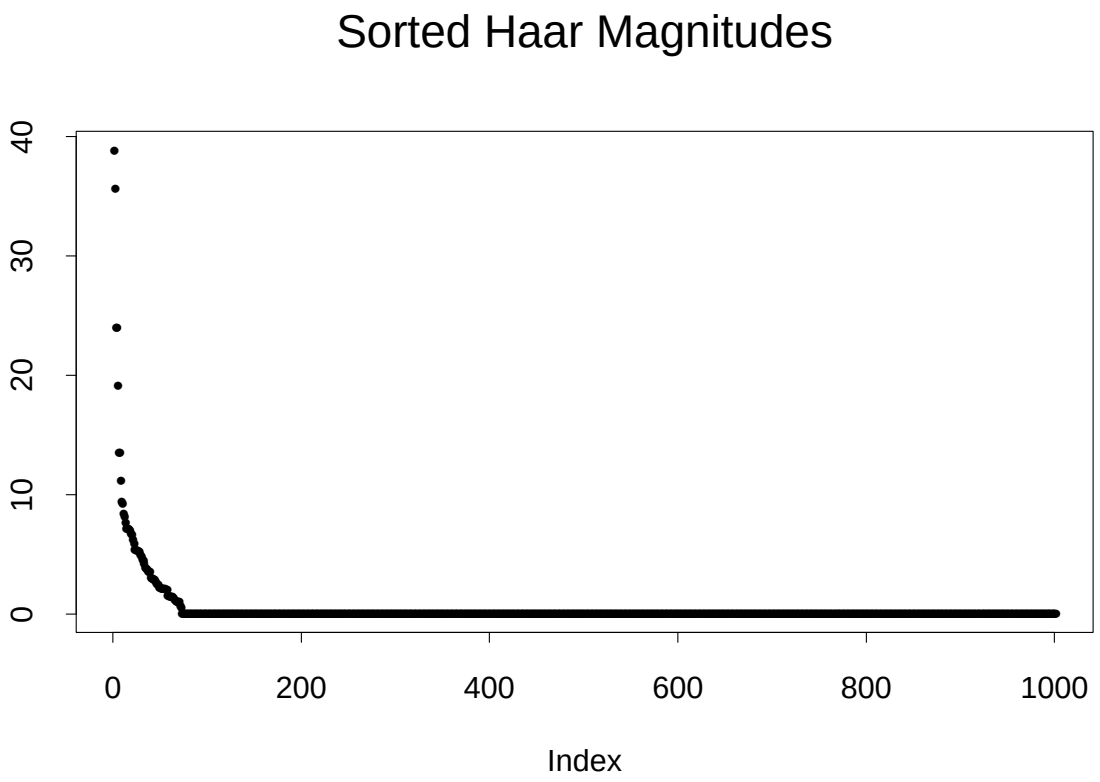


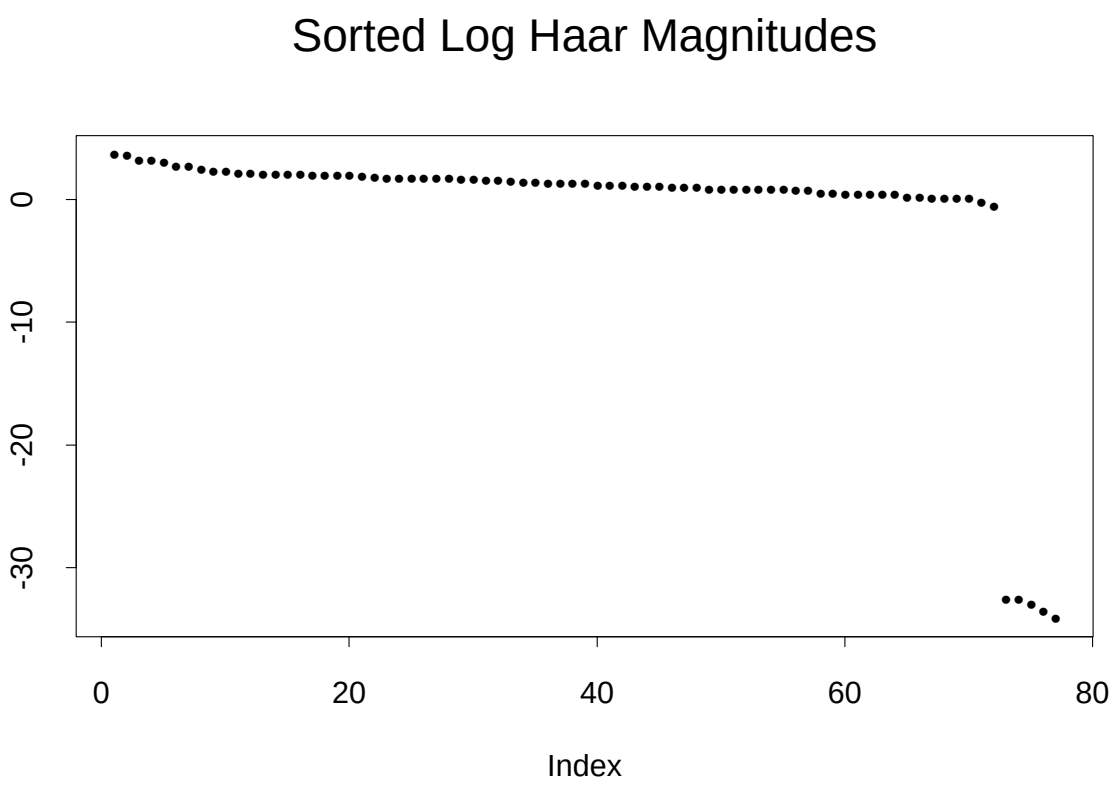
Fourier Signal Reconstruction 10%



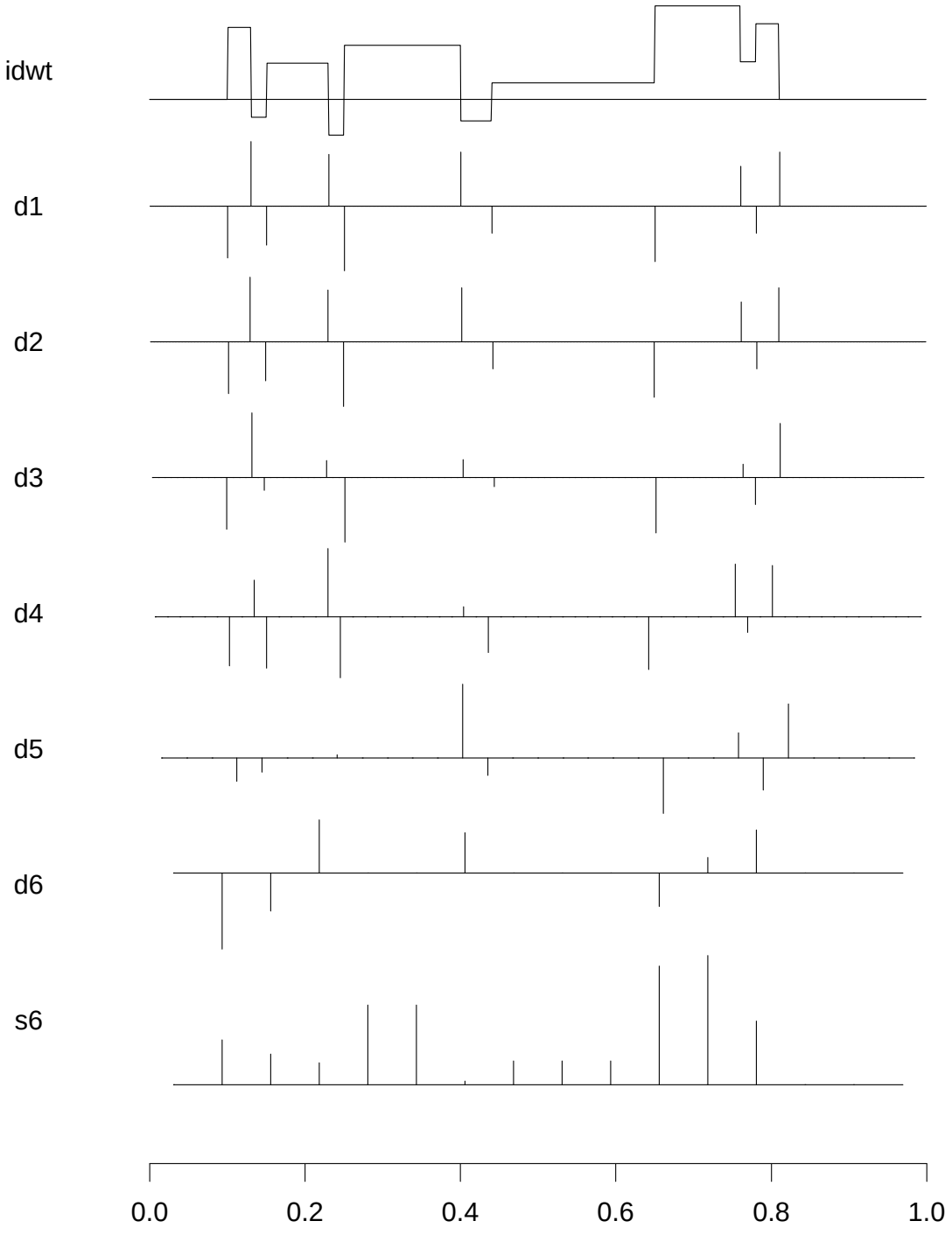
Fourier Signal Reconstruction 25%



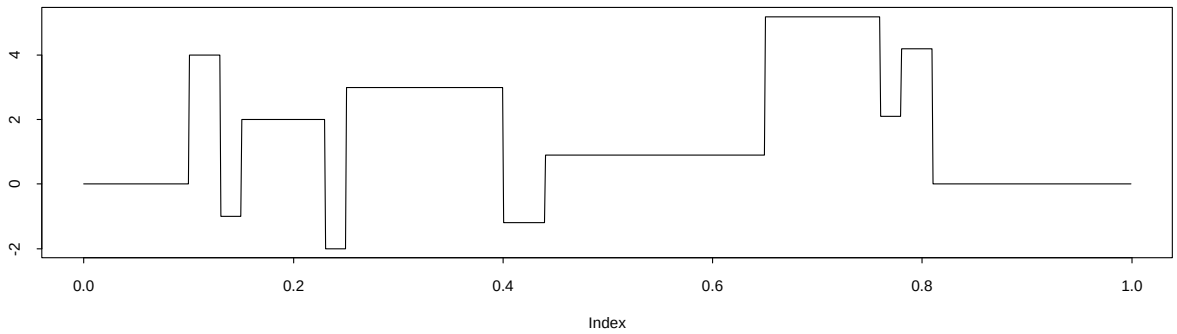

Sorted Haar Magnitudes
0
10
20
30
40
0
200
400
600
800
1000
Index


Sorted Log Haar Magnitudes
0
-10
-20
-30
0
20
40
60
80
Index

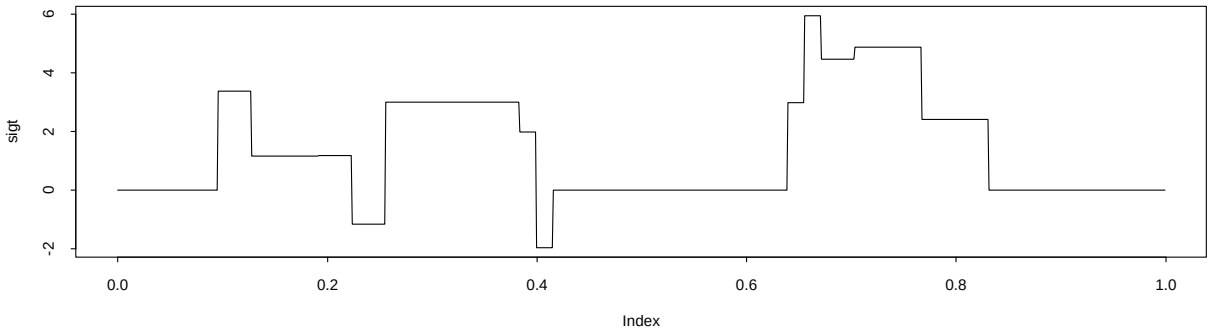
Wavelet Decomposition



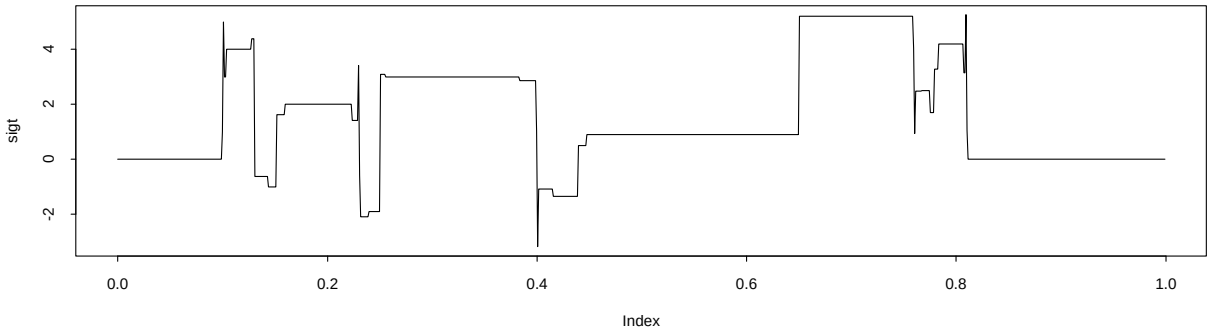
Signal



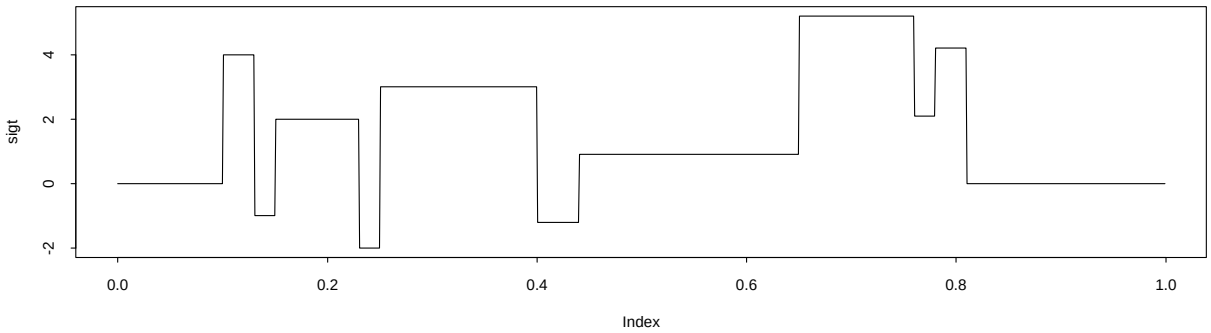
Wavelet Signal Reconstruction 1%

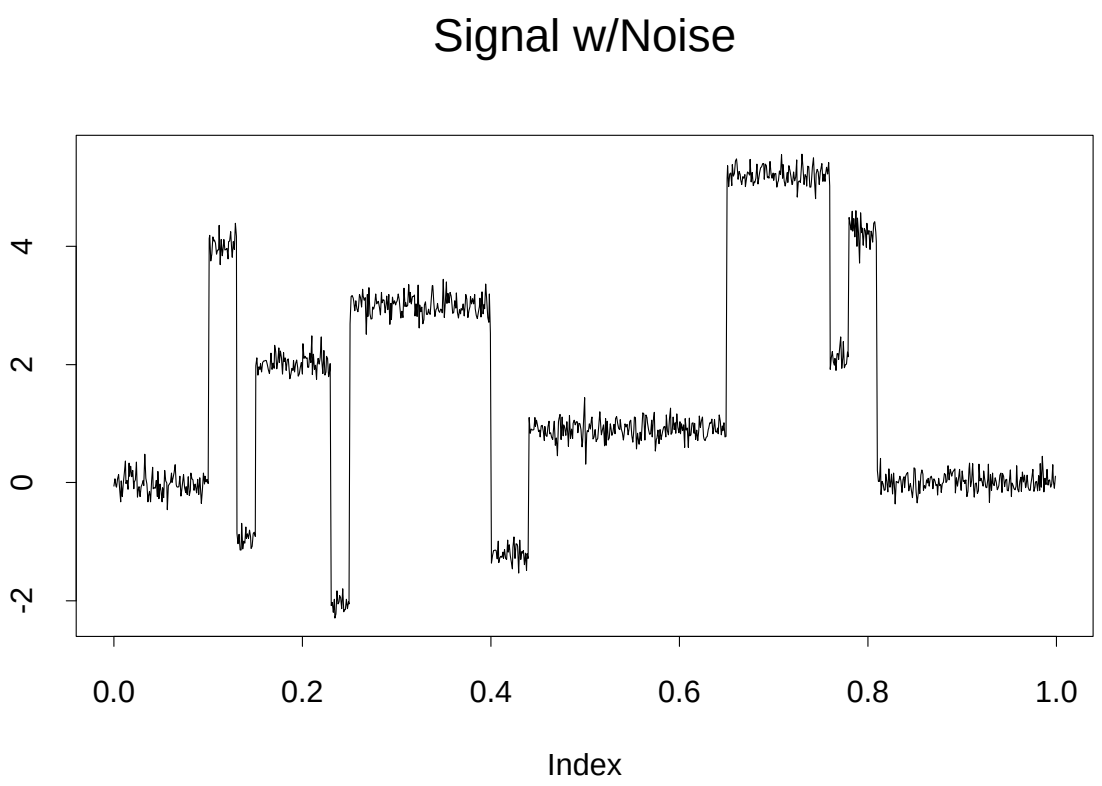


Wavelet Signal Reconstruction 5%

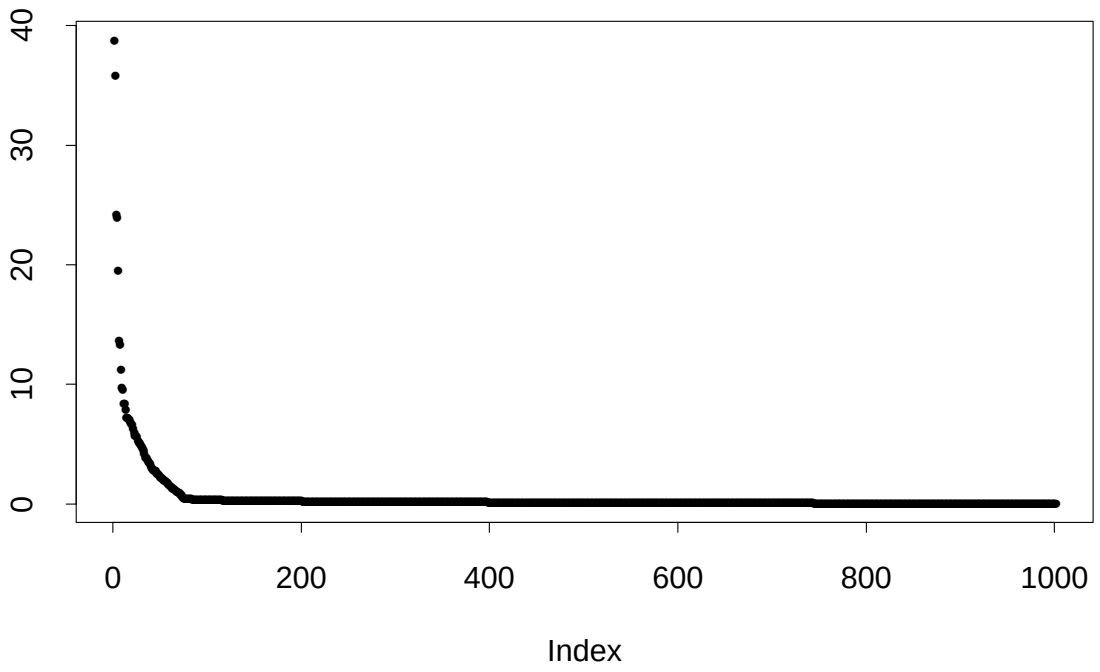


Wavelet Signal Reconstruction 8%

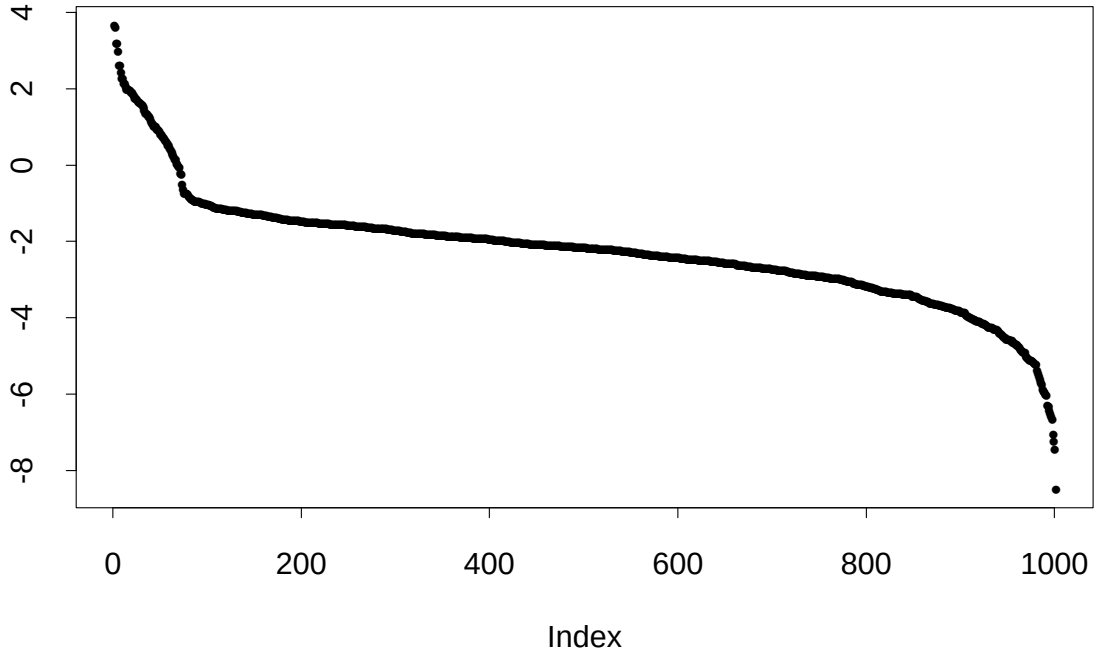



Signal w/Noise
4
2
0
-2
0.0
0.2
0.4
0.6
0.8
1.0
Index

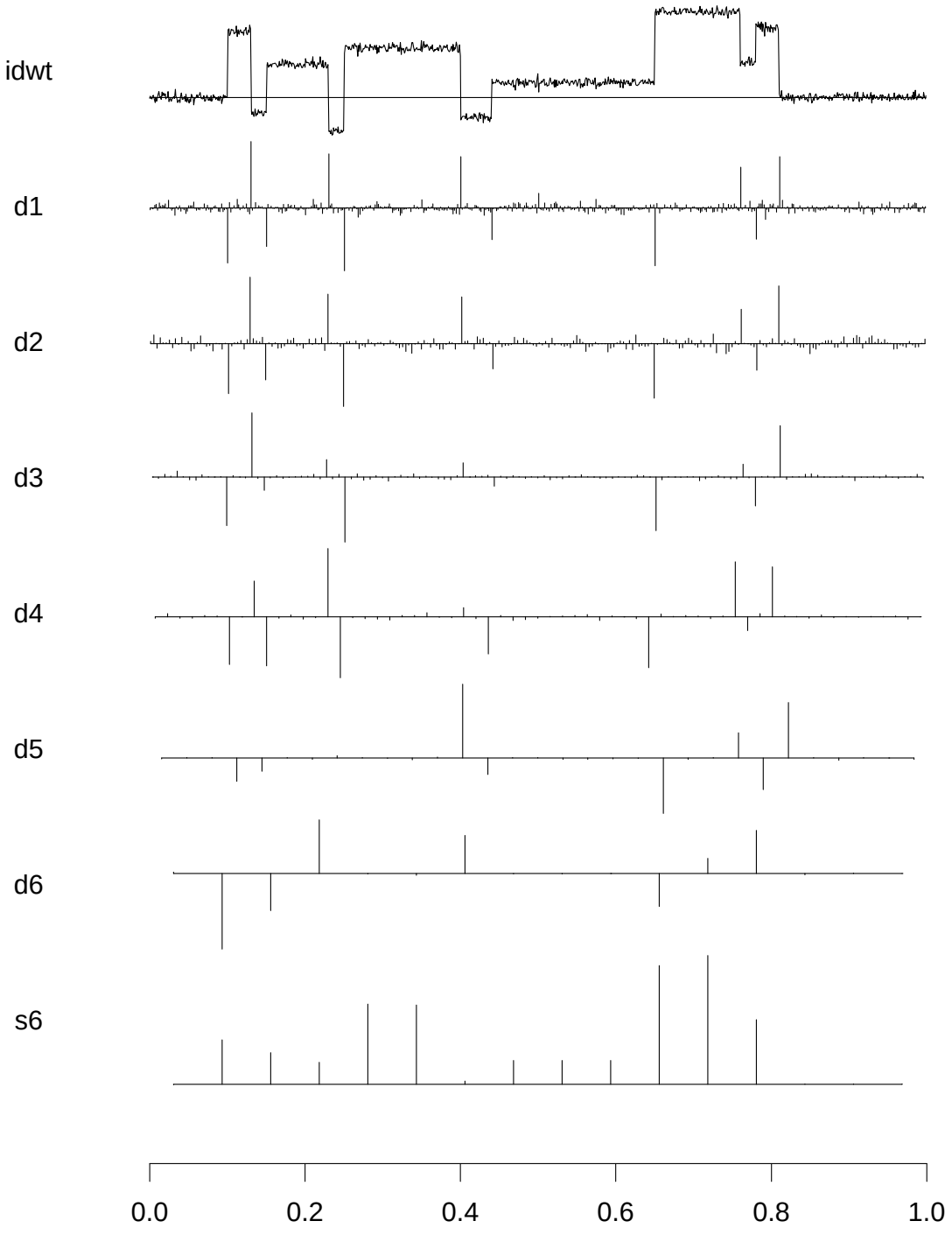
Signal w/Noise Sorted Haar Magnitudes



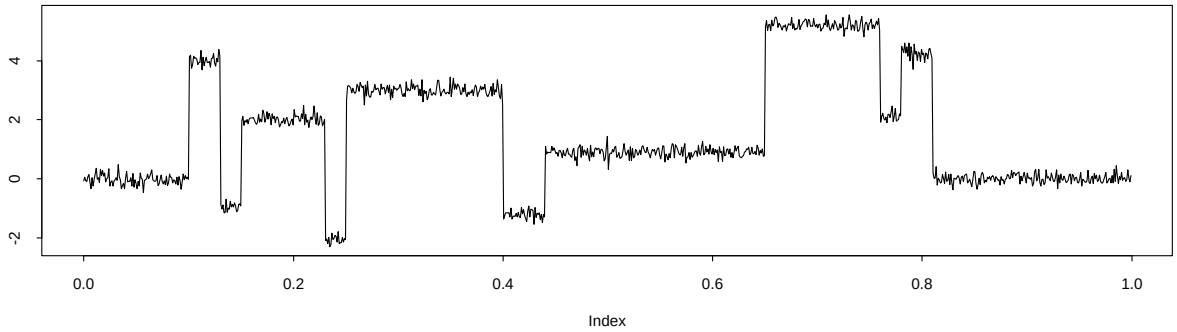
Signal w/Noise Sorted Log Haar Magnitudes



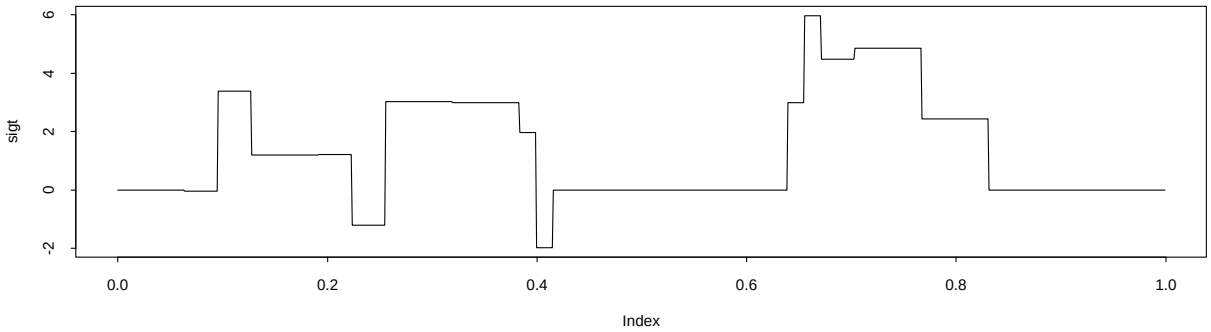
Signa; w/Noise Wavelet Decomposition



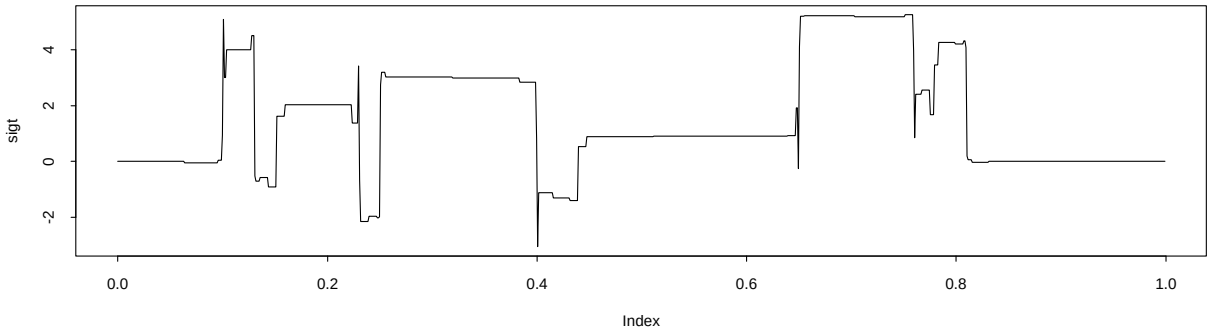
Signal w/Noise



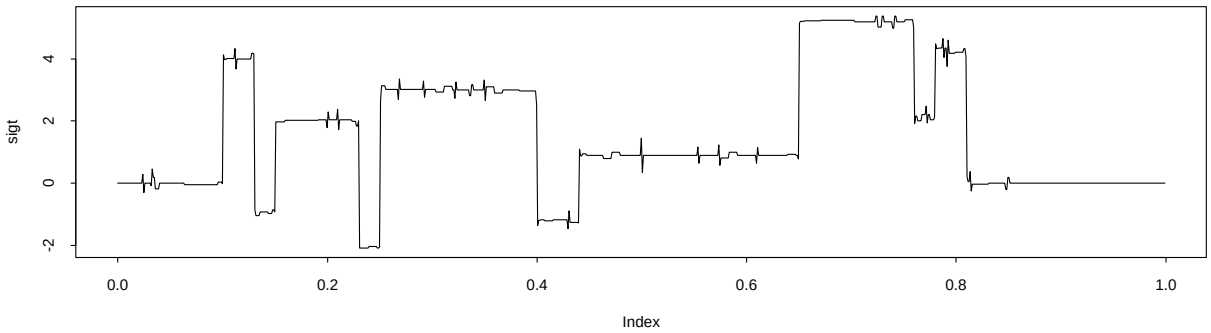
Wavelet Signal w/Noise Reconstruction 1%



Wavelet Signal w/Noise Reconstruction 5%



Wavelet Signal w/Noise Reconstruction 10%



Procedure

1. Compute the discrete wavelet transform.

2. Extract the fine-scale coefficients, d_{jk}.

3. Replace each d_{jk} with

$$\begin{array}{ll} 0 & \text{if } |d_{jk}| \leq \sigma \\ sign(d_{jk})(|d_{jk}| - \sigma) & \text{if } |d_{jk}| > \sigma \end{array}$$

4. Compute the inverse discrete wavelet transform.

σ based on truncation levels

5: Principal Component Analysis

a: Motivation

b: Definition

c: ChromatographyExample

Motivation

What is the best choice of basis for a set of data?

Design of best basis:

Let $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ be n repetitions of a p-dimensional vector. Let u_{ij} be the jth element of the ith vector.

$$\text{Ex: } \boldsymbol{u} = \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \\ u_{51} \\ u_{61} \end{bmatrix} = \begin{bmatrix} 5.1 \\ 5.8 \\ 5.9 \\ 5.4 \\ 6.0 \\ 5.3 \end{bmatrix}, \text{ etc.}$$

$$\boldsymbol{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pn} \end{bmatrix} \quad p \times n$$

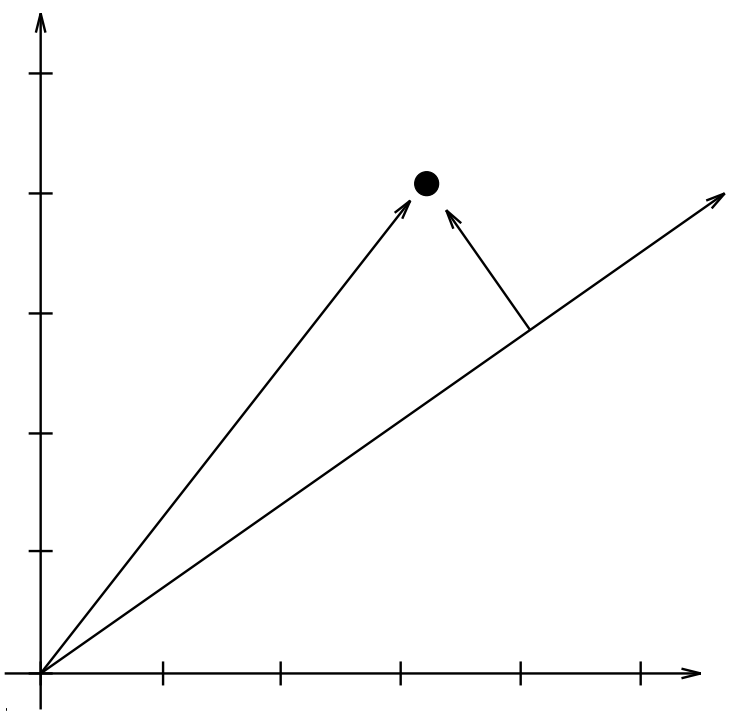
Motivation

Want to do the best job with only one basis vector $\boldsymbol{b}_1$.

Let c_i minimize $\|\boldsymbol{u}_i - c_i\boldsymbol{b}_1\|^2 = \sum_j (u_{ij} - c_i b_{1j})^2$

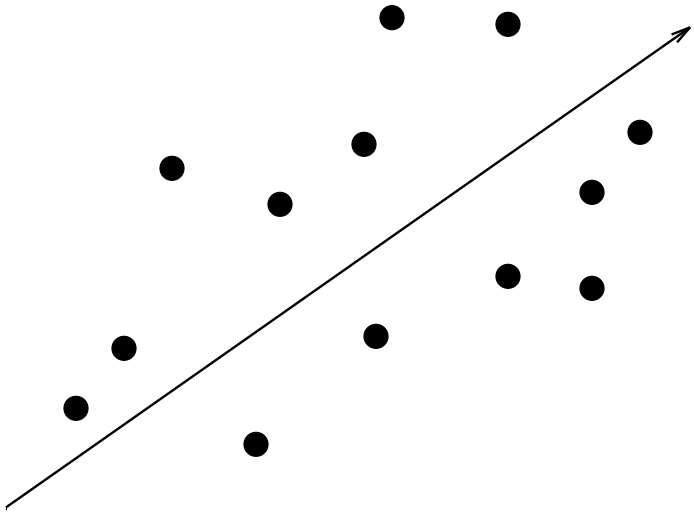
$c_i\boldsymbol{b}_1$ is the closest you can get to $\boldsymbol{u}_i$ with only basis vector $\boldsymbol{b}_1$. $(c_i = < \boldsymbol{b}_1, \boldsymbol{u}_i >)$

Want to find $\boldsymbol{b}_1$ to minimize
$\sum_i \|\boldsymbol{u}_i - c_i\boldsymbol{b}_1\|^2$

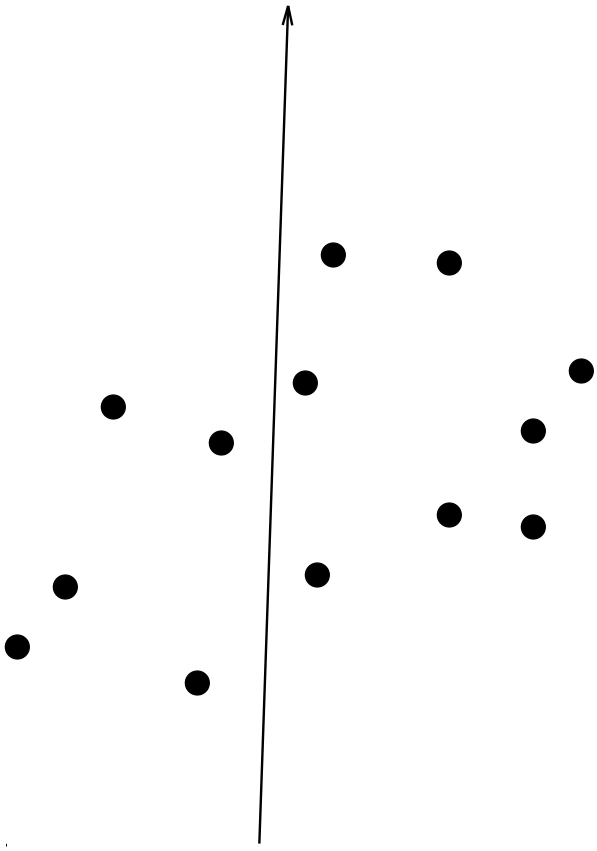


Motivation

Good Direction



Bad Direction



Motivation

Want to do the best job with only two basis vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$, where $< \boldsymbol{b}_1, \boldsymbol{b}_2 >= 0$.

Let $c_{1i} =< \boldsymbol{b}_1, \boldsymbol{u}_i >$ and $c_{2i} =< \boldsymbol{b}_2, \boldsymbol{u}_i >$.

Given $\boldsymbol{b}_1$, want to find $\boldsymbol{b}_2$ to minimize

$$\sum_i \|\boldsymbol{u}_i - (c_{1i}\boldsymbol{b}_1 + c_{2i}\boldsymbol{b}_2)\|^2$$

Continue scheme until you have a basis $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$, ordered from the most important vector to the least.

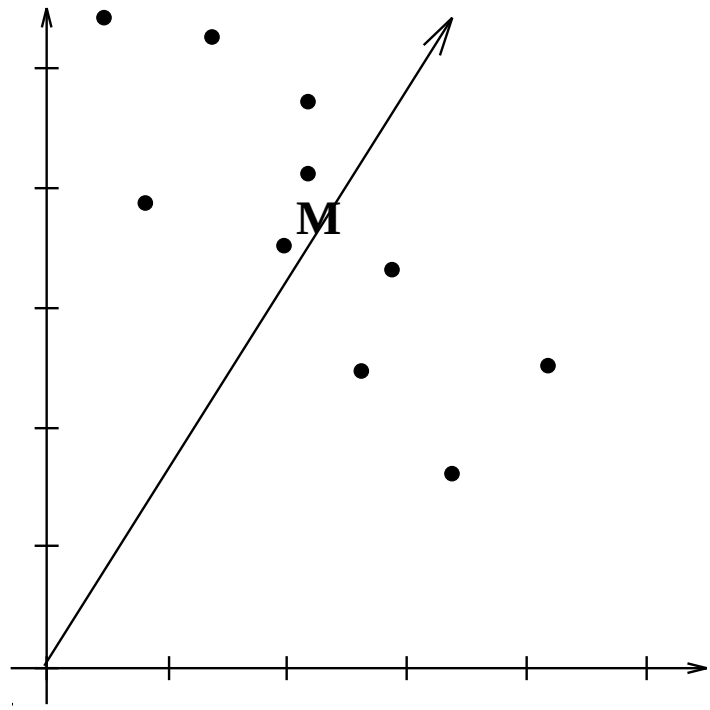
Effect of the Mean Vector

Given that the mean has been subtracted off, want to do the best job with only one basis vector $\boldsymbol{b}_1$.

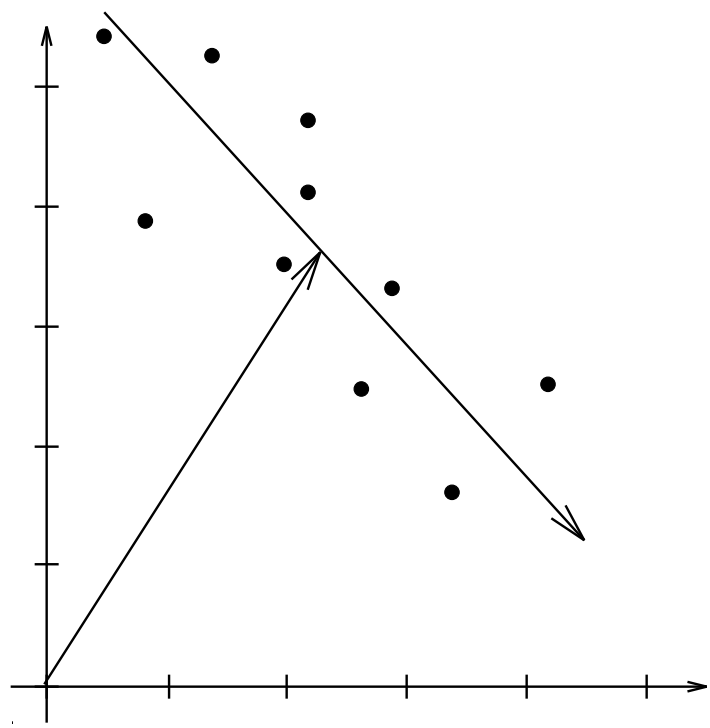
Want to find $\boldsymbol{b}_1$ to minimize
$\sum_i \|\boldsymbol{u}_i - (\boldsymbol{E}(\boldsymbol{u}) + c_i \boldsymbol{b}_1)\|^2$

Effect of the Mean Vector

Without Mean Vector



With Mean Vector

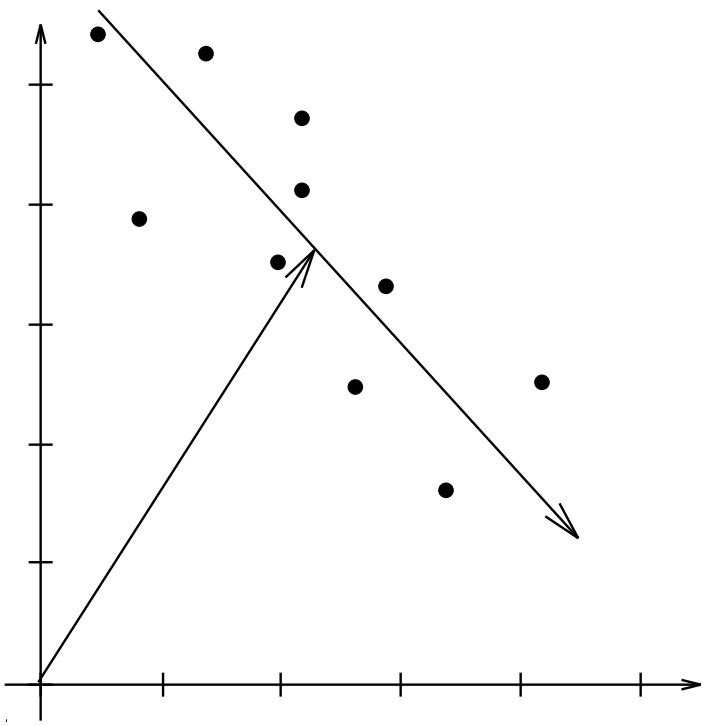


Classical Definition

Note:

(1) Minimizing residual^2 $\equiv$ maximizing length^2.

(2) Length with respect to the mean $\equiv$ variance.



Classical Definition

The *first principal component* is the linear combination (of the variables) $\boldsymbol{b}_1^T \boldsymbol{U}$ with maximal variance, such that $\|\boldsymbol{b}_1\| = 1$.

The *second principal component* is the linear combination (of the variables) $\boldsymbol{b}_2^T \boldsymbol{U}$ with maximal variance, such that $\|\boldsymbol{b}_2\| = 1$ and $< \boldsymbol{b}_1, \boldsymbol{b}_2 >= 0$.

$$\begin{bmatrix} - & \boldsymbol{b}_1 & - \\ - & \boldsymbol{b}_2 & - \\ \vdots & \vdots & \vdots \\ - & \boldsymbol{b}_n & - \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pn} \end{bmatrix}$$

$$= \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{p1} & v_{p2} & \cdots & v_{pn} \end{bmatrix} p \times n$$

$$\boldsymbol{BU} = \boldsymbol{V}$$

$$p \times p \quad p \times n = p \times n$$

Decomposition of Variance

Let σ_i^2 be the variance of the ith variable of $\boldsymbol{u}$.

Let ν_i^2 be the variance of the ith variable of $\boldsymbol{v}$ (The ith principal component).

Important Result: $\displaystyle\sum_{i=1}^{p} \sigma_i^2 = \sum_{i=1}^{p} \nu_i^2$.