Numerical Explorations for Fast Spectrum of Fractional Gaussian Noise

Martin Mächler
ETH Zurich
June 2011
(LATeX’ed February 5, 2020)

Abstract

The package longmemo ....
Paxson (1997) ...

Keywords: Euler-Maclaurin Formula, Fractional Gaussian Noise, Spectrum.

1. .. intro ..

The spectral density of fractional Gaussian noise (“fGn”) with Hurst parameter $H \in (0, 1)$ is (Beran (1986, 1994))

$$f_H(\lambda) = A(\lambda, H) \left(|\lambda|^{-2H-1} + B(\lambda, H)\right),$$

for $\lambda \in [-\pi, \pi]$, where $A(\lambda, H) = 2 \sin(\pi H) \Gamma(2H + 1) (1 - \cos \lambda)$, and

$$B(\lambda, H) = \sum_{j=1}^{\infty} \left( (2\pi j + \lambda)^{-(2H+1)} + (2\pi j - \lambda)^{-(2H+1)} \right).$$

For the Whittle estimator of $H$ and also other purposes, it’s advantageous to be able to evaluate $f_H(\lambda_i)$ efficiently for a whole vector of $\lambda_i$, typically Fourier frequencies $\lambda_i = 2\pi i / n$, for $i = 1, 2, \ldots, \lfloor (n-1)/2 \rfloor$. Such evaluation is problematic because of the infinite sum for $B(\lambda, H)$ in (2).

Traditionally, e.g., already in Appendix.... of Beran (1994), the infinite sum $\sum_{j=1}^{\infty}$ had been replaced by $\sum_{j=1}^{200} —$ which was still not very efficient and not extremely accurate. In our R package longmemo, we now provide the function $B\text{.specFGN}(\lambda, H)$ to compute $B(\lambda, H)$, using several ways to compute the infinite sum approximately, e.g., for $H = 0.75$ and $n = 500$, i.e., at 250 Fourier frequencies,

```r
> require("longmemo")
> fr <- .ffreq(500)
> B.1 <- B.specFGN(fr, H = 0.75, nsum = 200, k.approx=NA)
> B.xct <- B.specFGN(fr, H = 0.75, nsum = 10000, k.approx=NA)
> all.equal(B.xct, B.1)
[1] "Mean relative difference: 0.0001243095"
```

[1] "Mean relative difference: 0.0001243095"
which means that the 200 term approximation is accurate to 4 decimal digits for $H = .75$ but the accuracy is smaller for smaller $H$.

For this reason, Paxson (1997) derived formulas for fast and stilly quite accurate approximations of $B(\lambda, H)$, noting that $B(\lambda, H) = \sum_{j=1}^{\infty} f(j; \lambda, H)$ for

$$f(x; \lambda, H) = (2\pi x + \lambda)^{-(2H+1)} + (2\pi x - \lambda)^{-(2H+1)},$$

(3)

and the fact that $\sum_{j=1}^{\infty} f(j)$ is a Riemann sum approximation of $\int_{0}^{\infty} f(x) \, dx$ or $\int_{1}^{\infty} f(x) \, dx$.

```r
> fB <- function(x, lambda, H) {
    u <- 2 * pi * x
    h <- -(2 * H + 1)
    (u + lambda)^h + (u - lambda)^h
}
```

Now it’s clear that $f(x)$ cannot be computed (or “is infinite”) at $x = 0$, and more specifically, $f(x)$ tends to $\infty$ when $x \rightarrow \frac{\lambda}{2\pi}$, as in the second term of $f$, $2\pi x - \lambda$ only remains positive when $2\pi x > \lambda$. This is always fulfilled for $x \in \{1, 2, \ldots\}$, as $\lambda < \pi$, but is problematic when considering $\int_{0}^{b} f(x) \, dx$ as above. Some illustrations of the function $f(x; \lambda, H)$ and its “pole” at $\frac{\lambda}{2\pi}$:
So, very clearly, Paxson’s first formula, using \( f_0^1 f(x) \, dx \) is not feasible, as \( f(x) \) is not defined (or defined as \( \infty \)) for \( x \leq \lambda/(2\pi) \).

However, his generalized formula, “(7), p. 15”,

\[
\sum_{i=1}^{\infty} f_i \approx \sum_{j=1}^{k} f_j + \frac{1}{2} \int_{k}^{k+1} f(x) \, dx + \int_{k+1}^{\infty} f(x) \, dx,
\]

clearly is usable for \( k \geq 1 \) (but not for \( k = 0 \), contrary to what he suggests). Indeed, with \( \text{B.specFGN}(\lambda, H, k\text{.approx}) \), we now provide the result of applying approximation (4) to the infinite sum for \( B(\lambda, H) \) in (2).

Paxson ended the \( k = 3 \) approximation which he further improved considerably, empirically, by numerical comparison (and least squares fitting) with the “accurate” formula using \( \text{nsnum} = 10'000 \) terms. In the following section, we propose another improvement over Paxson’s original idea:

2. Better approximations using the Euler–MacLaurin formula

Copied straight from \( \text{http://en.wikipedia.org/wiki/Euler-Maclaurin_formula} \):

If \( n \) is a natural number and \( f(x) \) is a smooth, i.e., sufficiently often differentiable function defined for all real numbers \( x \) between 0 and \( n \), then the integral

\[
I = \int_{0}^{n} f(x) \, dx
\]
can be approximated by the sum (or vice versa)

\[ S = \frac{1}{2} f(0) + f(1) + \ldots + f(n - 1) + \frac{1}{2} f(n) \]

(see trapezoidal rule). The Euler–Maclaurin formula provides expressions for the difference between the sum and the integral in terms of the higher derivatives \( f^{(k)} \) at the end points of the interval 0 and n. Explicitly, for any natural number \( p \), we have

\[ S - I = \sum_{k=2}^{p} \frac{B_k}{k!} \left( f^{(k-1)}(n) - f^{(k-1)}(0) \right) + R \]

where \( B_1 = -1/2, B_2 = 1/6, B_3 = 0, B_4 = -1/30, B_6 = 1/42, B_8 = -1/30, \ldots \) are the Bernoulli numbers, and \( R \) is an error term which is normally small for suitable values of \( p \). (The formula is often written with the subscript taking only even values, since the odd Bernoulli numbers are zero except for \( B_1 \).)

Note that

\[ -B_1(f(n) + f(0)) = \frac{1}{2} (f(n) + f(0)). \]

Hence, we may also write the formula as follows:

\[ \sum_{i=0}^{n} f(i) = \int_{0}^{n} f(x) \, dx - B_1(f(n) + f(0)) + \sum_{k=1}^{p} \frac{B_{2k}}{(2k)!} \left( f^{(2k-1)}(n) - f^{(2k-1)}(0) \right) + R. \quad (6) \]

In the context of computing asymptotic expansions of sums and series, usually the most useful form of the Euler–Maclaurin formula is

\[ \sum_{n=a}^{b} f(n) \sim \int_{a}^{b} f(x) \, dx + \frac{f(a) + f(b)}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} \left( f^{(2k-1)}(b) - f^{(2k-1)}(a) \right), \]

where \( a \) and \( b \) are integers. Often the expansion remains valid even after taking the limits \( a \to -\infty \) or \( b \to +\infty \), or both. In many cases the integral on the right-hand side can be evaluated in closed form in terms of elementary functions even though the sum on the left-hand side cannot.

(end of citation from Wikipedia)
• Running under: Fedora 30 (Thirty)
• Matrix products: default
• BLAS: /u/maechler/R/D/r-patched/F30-64-inst/lib/libRblas.so
• LAPACK: /u/maechler/R/D/r-patched/F30-64-inst/lib/libRlapack.so
• Base packages: base, datasets, grDevices, graphics, methods, stats, utils
• Other packages: longmemo 1.1-2
• Loaded via a namespace (and not attached): compiler 3.6.2, sfsmisc 1.1-3, tools 3.6.2

4. Conclusion

References


Affiliation:

Martin Mächler
Seminar für Statistik, HG G 16
ETH Zurich
8092 Zurich, Switzerland
E-mail: maechler@stat.math.ethz.ch
URL: http://stat.ethz.ch/people/maechler