Package ‘wfGn’

October 29, 2010

Type Package

Title Estimation of the Hurst Parameter for Contaminated Fractional Gaussian Noises

Version 1.2-1

Date 2010-10-29

Author Wonsang You

Maintainer Wonsang You <you@ifn-magdeburg.de>

Description Estimation of Hurst parameter of a fractional Gaussian noise on the basis of the modified Whittle maximum likelihood estimator in presence of outliers or an additive noise

License GPL (>= 2.0)

LazyLoad yes

Depends multitaper, rgenoud, snow, Matching

R topics documented:

`wfGn-package` *Estimation of the Hurst Parameter for Contaminated Fractional Gaussian Noises*

Description

Estimates the Hurst parameter of a contaminated fractional Gaussian noises by using the modified Whittle estimator.

Details

Package:	wfGn
Type:	Package
Version:	1.2-1
Date:	2010-10-29
License:	GPL (>=2.0)
LazyLoad:	yes

Author(s)

Wonsang You and Jan Beran

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Percival and Walden (2000) Wavelet Methods for Time Series Analysis, Cambridge University Press.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

S. Achard and J.-F. Coeurjolly (2009). Discrete variations of the fractional Brownian in the presence of outliers and an additive noise.

Examples

```
n<-10000;H<-0.2
z<-perturbFGN(n,H,type="WN",SNR=-10,plot=TRUE)
wFGN(z)
```

`CetaFGN` *Covariance Matrix of fGn Parameters*

Description

Covariance matrix of parameters $\widehat{\eta}$ of fractional Gaussian noise.

Usage

```
CetaFGN(eta,snr=NULL)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`snr`	the signal-to-noise ratio. Default is `NULL`.

Details

The most parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran. See Beran (1994) for details.

Value

A covariance matrix of the estimated fGn parameters with size $M \times M$ where M is the length of $\widehat{\eta}$.

Author(s)

Jan Beran (original) and Wonsang You (fine tuning)

References

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

fspecFGN

Examples

```
eta <- c(H=0.7); SNR <- 10
cov <- CetaFGN(eta,SNR)
```

`circFGN` *Simulation of a fractional Gaussian noise*

Description

Generates a sample path of fractional Gaussian noise with scaling coefficient `C=1` and Hurst parameter `H` in (0,1)

Usage

```
circFGN(n, H, C=1, plot=FALSE)
```

Arguments

`n`	sample size
`H`	Hurst parameter
`C`	scaling coefficient
`plot`	plot of the generated sample path

Details

This simulation is achieved by computing the lag difference of a fractional Brownian motion which is generated by the method in the package `longmemo`.

Value

a vector of length `n`.

Author(s)

Jan Beran (original) and Wonsang You (fine tuning)

References

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`perturbFGN`, `wFGN`

Examples

```
n<-10000; H<-0.2
ts <- circFGN(n,H)
```

`Csum` *Cumulative Qeta Difference Function*

Description

Cumulative Qeta Difference Function to be minimized to estimate the parameters of a fractional Gaussian noise.

Usage

```
Csum(eta,n,yper,snr,pertype)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`n`	the number of time points.
`yper`	a vector of periodogram with length of the largest integer less than `(n-1)/2`
`snr`	the signal-to-noise ratio.
`pertype`	the type of periodogram. Possible modes are `"per"`,`"taper"`.

Details

Let $I(f_i)$ and $S(f_i)$ be respectively the periodogram of a given perturbed fractional Gaussian noise and the spectral density of perturbed fGn with Hurst exponent `eta` and the signal-to-noise ratio `SNR` where $f_i = 2\pi * i/n$ with $i = 1, ..., (n-1)/2$. Then, the value of cumulative Qeta difference function is determined as
$\eta = \frac{1}{m^*} \sum_{t=1}^{m^*} (I_t - S_t)^2$
where
$I_t = \sum_{j=1}^{t} I(f_j) / \sum_{j=1}^{m^*} I(f_j)$ and $S_t = \sum_{j=1}^{t} S_X(f_j) / \sum_{j=1}^{m^*} S_X(f_j)$.
Some parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran. See Beran (1994) for details.

Value

`A,B,Tn`	defined in Qeta.
`z`	the test statistics
`pval`	the p-value
`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.
`value`	a value for minimization

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`Qeta`, `Lpvar`, `QLCfun`

Examples

```
n<-1000; H<-0.7; SNR<-10
ts <- perturbFGN(n,H,type="WN",SNR=SNR)
yper<-per(ts)
ts.csum<-Csum(H,n,yper,snr=SNR,pertype="per")
```

`fspecBFGN` *Spectral Density of a bivariate Fractional Gaussian Noise*

Description

Generation of the spectral density of a bivariate fractional Gaussian noise.

Usage

```
fspecBFGN(eta1, eta2, m)
```

Arguments

`eta1`	the Hurst exponent of the first component of a bivariate fractional Gaussian noise.
`eta2`	the Hurst exponent of the second component of a bivariate fractional Gaussian noise.
`m`	the number of time points.

Details

This function produces the spectral density of a bivariate fractional Gaussian noise.

Value

`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`fspecFGN`, `fspecPFGN`, `wFGN`

Examples

```
eta1<-0.7; eta2<-0.8; m<-1000
fspec<-fspecBFGN(eta1,eta2,m)
```

`fspecFGN` *Spectral Density of a Fractional Gaussian Noise*

Description

Generation of the spectral density of a fractional Gaussian noise.

Usage

```
fspecFGN(eta,m)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`m`	the number of time points.

Details

This function produces the spectral density of a fractional Gaussian noise.

Value

`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.

Author(s)

Jan Beran (original) and Wonsang You (fine tuning)

References

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

fspecBFGN, fspecPFGN, wFGN

Examples

```
eta<-0.7; m<-1000
fspec<-fspecFGN(eta,m)
```

---

fspecPFGN — *Spectral Density of a perturbed Fractional Gaussian Noise*

Description

Generation of the spectral density of a fractional Gaussian noise contaminated by white noise.

Usage

```
fspecPFGN(eta, m, SNR=NULL)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`m`	the number of time points.
`SNR`	the signal-to-noise ratio.

Details

This function produces the spectral density of a fractional Gaussian noise contaminated by white noise.

Value

`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`fspecFGN`, `fspecPFGN`, `wFGN`

Examples

```
eta<-0.7; m<-1000; SNR=0
fspec<-fspecPFGN(eta,m,SNR)
```

`Lpvar` *Log Periodogram Variance Function*

Description

Log Periodogram Variance Function to be minimized to estimate the parameters of a fractional Gaussian noise.

Usage

```
Lpvar(eta,n,yper,snr,pertype)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`n`	the number of time points.
`yper`	a vector of periodogram with length of the largest integer less than `(n-1)/2`
`snr`	the signal-to-noise ratio.
`pertype`	the type of periodogram. Possible modes are `"per"`,`"taper"`.

Details

Let $I(f_i)$ and $S(f_i)$ be respectively the periodogram of a given perturbed fractional Gaussian noise and the spectral density of perturbed fGn with Hurst exponent `eta` and the signal-to-noise ratio `SNR` where $f_i = 2\pi * i/n$ with $i = 1, ..., (n-1)/2$. Then, the value of log periodogram variance function is determined as

$\epsilon(f) \equiv \log\left(\frac{I(f)}{S_X(f)}\right) + \gamma.$

See Percival and Walden (2000) for details. Some parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran.

Value

`A,B,Tn`	defined in Qeta.
`z`	the test statistics
`pval`	the p-value
`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.
`value`	a value for minimization

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Percival and Walden (2000) Wavelet Methods for Time Series Analysis, Cambridge University Press.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

Qeta, Csum, QLCfun

Examples

```
n<-1000; H<-0.7; SNR<-10
ts <- perturbFGN(n,H,type="WN",SNR=SNR)
yper<-per(ts)
ts.lpvar<-Lpvar(H,n,yper,snr=SNR,pertype="per")
```

`per` *Periodogram Estimator*

Description

Estimation of periodogram of a given time series.

Usage

```
per(z)
```

Arguments

`z`	a vector which indicates a time series

Value

a vector of periodogram with length of $1 + floor(n/2)$ where n is the length of z. This function is originated from Jan Beran. See Beran (1994) for details.

Author(s)

Jan Beran

References

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

Examples

```
n<-1000; H<-0.7;
ts <- circFGN(n,H)
yper<-per(ts)
```

---

perturbFGN *Simulation of a perturbed fGn*

Description

Simulation of a sample path of a fractional Gaussian noise contaminated by outliers or an additive Gaussian noise

Usage

```
perturbFGN(n, H, C = 1, type = "no", p=.005, SNR=NULL, plot = FALSE)
```

Arguments

`n`	sample size
`H`	Hurst parameter in (0,1)
`C`	scaling coefficient. Default is `1`.
`type`	type of perturbation. Possible modes are `"no"`,`"WN"`,`"AO"`.
`p`	the ratio of perturbation by outliers. Default is 0.005.
`SNR`	Signal to noise ratio
`plot`	plot of perturbed fGn. Default is `FALSE`.

Details

The types of contamination are listed below. The most parts of this function were adopted from the package `dvfBm`.

"no" no contamination

"AO" contamination by additive outliers of 0.5%

"WN" contamination by i.i.d. standard Gaussian noise

Value

a vector of length `n`.

Author(s)

Wonsang You

References

S. Achard and J.-F. Coeurjolly (2009). Discrete variations of the fractional Brownian in the presence of outliers and an additive noise.

See Also

`circFGN`, `wFGN`

Examples

```
ts1<-perturbFGN(1000,0.2,type="no",plot=TRUE)
ts2<-perturbFGN(1000,0.7,type="AO",SNR=20,plot=TRUE)
ts3<-perturbFGN(10000,0.3,type="WN",SNR=10,plot=TRUE)
```

`Qeta` *Qeta function*

Description

Function to be minimized to obtain the maximum likelihood of a fractional Gaussian noise.

Usage

```
Qeta(eta,n,yper,snr,pertype)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`n`	the number of time points.
`yper`	a vector of periodogram with length of the largest integer less than `(n-1)/2`
`snr`	the signal-to-noise ratio.
`pertype`	the type of periodogram. Possible modes are `"per"`,`"taper"`.

Details

Let $I(f_i)$ and $S(f_i)$ be respectively the periodogram of a given perturbed fractional Gaussian noise and the spectral density of perturbed fGn with Hurst exponent `eta` and the signal-to-noise ratio `SNR` where $f_i = 2\pi * i/n$ with $i = 1, ..., (n-1)/2$. Then, the value is determined as

$B = \frac{2\pi}{n} \sum_i 2 \times \left[\frac{I(f_i)}{S(f_i)}\right]^2$.

Also, A and Tn are defined as follows.

$A = \frac{2\pi}{n} \sum_i 2 \times \left[\frac{I(f_i)}{S(f_i)}\right]$, $T_n = \frac{A}{B^2}$

The most parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran. See Beran (1994) for details.

Value

`A,B,Tn`	defined in the above section
`z`	the test statistics
`pval`	the p-value
`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.
`value`	a value for minimization

Author(s)

Jan Beran (original) and Wonsang You (modifying)

References

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`Csum`, `Lpvar`, `QLCfun`

Examples

```
n<-1000; H<-0.7; SNR<-10
ts <- perturbFGN(n,H,type="WN",SNR=SNR)
yper<-per(ts)
ts.qeta<-Qeta(H,n,yper,snr=SNR,pertype="per")
```

`QLCfun` *Combinational Minimization Function*

Description

Combinational function of Qeta, Cumulative Qeta Difference, and Log Periodogram Variance Function to be minimized to estimate the parameters of a fractional Gaussian noise.

Usage

```
QLCfun(eta,n,yper,snr,pertype,weights)
```

Arguments

`eta`	a positive value of the Hurst exponent which is less than 1.
`n`	the number of time points.
`yper`	a vector of periodogram with length of the largest integer less than `(n-1)/2`
`snr`	the signal-to-noise ratio.
`pertype`	the type of periodogram. Possible modes are `"per"`,`"taper"`.
`weights`	a vector of length 3 which indicates the weights of Qeta, Cumulative Qeta Difference, and Log Periodogram Variance Functions

Details

Let $\widetilde{L}_1(y;\theta)$, $\widetilde{L}_2(y;\theta)$, $\widetilde{L}_3(y;\theta)$ be Qeta, Cumulative Qeta Difference, and Log Periodogram Variance Functions respectively. Also, let $p = (p_1, p_2, p_3)$ be a weight vector, and let $\widetilde{L}\,(y;\theta)$ be a combinational minimization function. Then, the value of the combinational function is determined as

$\widetilde{L}\,(y;\theta) = p_1 \times \widetilde{L}_1\,(y;\theta) + p_2 \times \widetilde{L}_2\,(y;\theta) + p_3 \times \widetilde{L}_3\,(y;\theta)$.

Some parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran. See Beran (1994) for details.

Value

`A,B,Tn`	defined in Qeta.
`z`	the test statistics
`pval`	the p-value
`fspec`	a vector of spectral density with length of the largest integer less than `(m-1)/2`.
`theta1`	a value of the first component of theta.
`value`	a value for minimization

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Percival and Walden (2000) Wavelet Methods for Time Series Analysis, Cambridge University Press.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

Qeta, Csum, Lpvar

Examples

```
n<-1000; H<-0.7; SNR<-10
ts <- perturbFGN(n,H,type="WN",SNR=SNR)
yper<-per(ts)
ts.qlc<-QLCfun(H,n,yper,snr=SNR,pertype="per")
```

---

wFGN *The Whittle estimator for a contaminated fGn*

Description

Whittle estimator of the Hurst parameter of a fractional Gaussian noise contaminated by additive outliers or noise.

Usage

```
wFGN(x, istart=1, iend=length(x), nloop=1, init=c(0.55,0.01),
ndeps=c(1e-7,1e-2), noise=TRUE, pertype="per", minfun="qeta",
weights=c(1,1,0), cluster=FALSE, print.level=2)
```

Arguments

`x`	a time series
`istart`	the index of start. Default is `1`
`iend`	the index of end. Default is the length of `x`.
`nloop`	the number of sub-series. The Hurst parameter of each sub-series is estimated respectively. Default is `1`.
`init`	the initial search values of Hurst parameter and SNR. Default is `c(0.55,0.01)`.
`ndeps`	A vector of step sizes for optimization. Default is `c(1e-7,1e-2)`.
`noise`	Enable the assumption of noise corruption. Default is `TRUE`.
`pertype`	type of periodogram. Possible modes are `"per"`, `"taper"`.

`minfun`	type of minimization function. Possible modes are `"qeta"`, `"lp"`, `"csum"`, and `"combi"`.
`weights`	A vector of weights for each minimization function when `minfun="combi"`.
`cluster`	A vector of machine names for parallel processing. For details, refer to the manual of package `genoud`.
`print.level`	The logic value which controls the printing level. Four levels are available: 0 (minimal), 1 (normal), 2 (detailed), and 3 (debug). Default is 2. Refer to the manual of package `rgenoud` for details.

Details

The Hurst parameter of a pure or corrupted fractional Gaussian noise is estimated by a modified Whittle estimator. The most parts of this function were adopted from the S-PLUS codes originally developed by Jan Beran. See Beran (1994) for details.

Value

`thetavector`	a vector of scaling coefficient and Hurst parameter
`Hlow`	the lower limit of 95% confidence interval of Hurst parameter
`Hup`	the upper limit of 95% confidence interval of Hurst parameter
`SNR`	the signal-to-noise ratio
`fest`	the spectral density of an estimated fGn

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

Jan Beran (1994) Statistics for Long-Memory Processes, Chapman & Hall.

See Also

`circFGN`, `perturbFGN`, `wFGN.eval`

Examples

```
n<-1000; H1<-0.2; H2<-0.7
ts1 <- circFGN(n,H1)
dat1 <- wFGN(ts1,noise=FALSE)
ts2 <- perturbFGN(n,H2,type="WN",SNR=10)
dat2 <- wFGN(ts2,pertype="per",minfun="combi")
```

`wFGN.eval` *Evaluation of the modified Whittle estimator for a contaminated fGn*

Description

Statistical performance evaluation of the modified Whittle estimator of the Hurst parameter of a fractional Gaussian noise contaminated by additive outliers or noise.

Usage

```
wFGN.eval(H=NULL, n=1000, m=100, type="no", SNR=NULL,
ndeps=c(1e-7,1e-2), noise=TRUE, pertype="per",
minfun="qeta", weights=c(1,1,0), cluster=FALSE,
plot=TRUE, sav=FALSE)
```

Arguments

`H`	Hurst parameter. If `H=NULL`, $H = 0.1, 0.2, ..., 0.9$ are tested. Default is `NULL`.
`n`	sample size. Default is `1000`.
`m`	the number of repetitions. Default is `100`.
`type`	type of perturbation. Possible modes are `"no"`,`"WN"`,`"AO"`. Refer to wFGN for details.
`SNR`	Signal to noise ratio.
`ndeps`	A vector of step sizes for optimization. Default is `c(1e-7,1e-2)`.
`noise`	Enable the assumption of noise corruption. Default is `TRUE`.
`pertype`	type of periodogram. Possible modes are `"per"`,`"taper"`.
`minfun`	type of minimization function. Possible modes are `"qeta"`,`"lp"`,`"csum"`,`"combi"`.
`weights`	A vector of weights for each minimization function when `minfun="combi"`.
`cluster`	A vector of machine names for parallel processing. For details, refer to the manual of package `genoud`.
`plot`	a boxplot of parameter estimation. Default is `TRUE`.
`sav`	Enable sample plots. Default is `FALSE`.

Details

The Hurst parameter of a fractional Gaussian noise is estimated by the modified Whittle estimator. This function evaluates the consistency of the Whittle estimator by several repetitions.

Value

Hdata	a $m \times 1$ or $m \times 9$ matrix of Hurst parameter estimates for fGn with different Hurst parameters
Hstat	a 4×1 or 4×9 matrix with a sample Hurst parameter, mean, standard deviation, and mean squared error (MSE) of Hurst parameter estimates
SNRdata	a $m \times 1$ or $m \times 9$ matrix of SNR estimates for fGn with different Hurst parameters
SNRstat	a 4×1 or 4×9 matrix with a sample SNR, mean, standard deviation, and mean squared error (MSE) of SNR estimates
Theta	a $m \times 1$ or $m \times 9$ matrix of scaling coefficient estimates

Author(s)

Wonsang You

References

Wonsang You (2010) Modified Whittle's Maximum Likelihood Estimator for Fractional Gaussian Noises Contaminated by Additive Noises, Technical Reports of the Leibniz Institute for Neurobiology, TR10015.

See Also

circFGN, perturbFGN, wFGN

Examples

```
dat <- wFGN.eval(H=0.2, n=10000, m=10, type="WN", SNR=-20, plot=TRUE)
```

Index