`package.qn` *Tools for robust empirical variogram estimation*

Description

qn is a collections of functions for robust variogram estimation.

Details

The first version was written for `S+` with the module `S+SpatialStats`. There was a need to port the functions to R.

The function `variogram.qn` is a literal translation and uses the same class construction as `variogram`.

Out of the many existing R packages to estimate the empirical variogram, I have chosen `geoR` and the function class of `variog.qn` is similar to `variog`. The `plot` method of `geoR` can be used.

Note

DISCLAIMER:

This is software for statistical research and not for commercial uses. The author does not guarantee the correctness of any function or program in this package. Any changes to the software should not be made without the author's permission.

References

Kaluzny, S. P. et al. (1996). S+SpatialStats - User's Manual for Windows and Unix, MathSoft, Inc., Seattle.

Rousseeuw, P.J. and Croux, C. (1993). "Alternatives to the Median Absolute Deviation," Journal of the American Statistical Association, Vol. 88, 1273-1283.

Genton, M. G., (1998). "Highly Robust Variogram Estimation", Mathematical Geology, Vol. 30, No. 2, 213-221.

Genton, M. G., (1998). "Spatial Breakdown Point of Variogram Estimators", Mathematical Geology, Vol. 30, No. 7, 853-871.

Furrer, R. and Genton, M. G. (1999). "Robust Spatial Data Analysis of Lake Geneva Sediments with S+SpatialStats", International Journal of Systems Research and Information Science, special issue on Spatial Data Analysis and Modeling, 8, 257-272.

qn-internal — *qn internal and secondary functions*

Description

Listed below are supporting functions for the package qn.

Usage

```
version.qn(verbose=TRUE)
```

qn — *Scale Estimator*

Description

Computes the highly robust estimator of scale 'qn'.

Usage

```
qn(x)
```

Arguments

x	real numeric matrix or vector. Missing or infinite values are filtered. (maximum length of x is 100000 elements).

Details

The estimator has a high breakdown point and a bounded influence function. The implementation given here is very fast (running in O(nlogn) time) and needs little storage space.

Value

estimator of scale 'qn'.

References

Rousseeuw, P.J. and Croux, C. (1993). "Alternatives to the Median Absolute Deviation," Journal of the American Statistical Association, Vol. 88, 1273-1283.

Examples

```
x <- c(rnorm(100,sd=4),rnorm(10,sd=100))
qn(x)
sd(x)
mad(x)
```

`variog.qn` *Empirical Variogram Estimation*

Description

Estimates the (semi)variogram for two dimensional spatial data using the highly robust estimator of scale `qn`.

Usage

```
variog.qn(geodata, coords = geodata$coords, data = geodata$data,
          uvec = "default",  max.dist, pairs.min = 2,
          messages.screen = TRUE, ...)
```

Arguments

`geodata`	a list containing elements `coords` and `data` as described next. Typically an object of the class `geodata` - a `geoR` data-set. If not provided the arguments `coords` must be provided instead.
`coords`	an n x 2 matrix containing coordinates of the n data locations in each row. Defaults to `geodata$coords`, if provided.
`data`	a vector or matrix with data values. If a matrix is provided, each column is regarded as one variable or realization. Defaults to `geodata$data`, if provided.
`uvec`	a vector with values defining the variogram binning. See `variog` for more details.
`max.dist`	a numerical value defining the maximum distance for the variogram. Pairs of locations separated for distance larger than this value are ignored for the variogram calculation. If not provided defaults takes the maximum distance among all pairs of data locations.
`pairs.min`	a integer number defining the minimum numbers of pairs for the bins.
`messages.screen`	logical. Indicates whether status messages should be printed on the screen (or output device) while the function is running.
`...`	Used only for simplicity that arguments of `variog` do not have to be removed

(Argument description has been taken from the `geoR` help).

Details

Since I omit any discussion about the variogram here, recommend reading the help of `variog` from the `geoR` package.

Method for class `variogram` include `plot`.

Value

An object of the `class variogram` which is a list with at least the following components:

`u`	a vector with distances.
`v`	a vector with estimated (semi)variogram values at distances given in `u`.
`n`	number of pairs in each bin.
`max.dist`	maximum distance between pairs allowed in the variogram calculations.
`uvec`	lags provided in the function call.
`n.data`	number of data.
`call`	the function call.
`output.type`	`bin`
`estimator.type`	`qn`
`trend`	`cte`
`direction`	`omnidirectional`

See Also

`qn`, `variog` from the `geoR` package.

Examples

```
library(geoR)

data(s100)
vario <- variog(s100, max.dist=1)
vario.qn <- variog.qn(s100, max.dist=1)

plot(vario.qn)
points(vario$u,vario$v,col=4)
```

`variogram.qn` *Empirical Variogram*

Description

Estimates the (semi)variogram for two dimensional spatial data using the highly robust estimator of scale `qn`.

Usage

```
variogram.qn(data, lag.vect, tol.lag = 0, minpairs = 6, printit = FALSE)
```

Arguments

`data`	a data frame containing the 3 columns. The first two form the locations while the third one is the numeric response. All variables must be vectors of equal length with no missing values (NAs).
`lag.vect`	a numeric vector containing the values of the distances at which the variogram is calculated.
`tol.lag`	lag tolerance.
`minpairs`	the minimum number of pairs of points (minimum value for np) that must be used in calculating a value. If np is less than minpairs then that value is dropped from the variogram.
`printit`	logical flag: the results are not prompted if FALSE (the default).

Details

The results are printed on local console if printit is TRUE.

The functions have been written to be used with the S+SpatialStats module of S+. I recommend to use `variog.qn` with `geoR`.

Value

an object of class `variogram` that inherits from `data.frame` with the following columns:

`distance`	the average distance for pairs in the lag.
`gamma`	the semi-variogram estimate.
`np`	the number of pairs in each lag.
`azimuth`	a factor denoting the angular direction zero.

The return object has an attribute call with an image of the call that produced the object.

See Also

`qn`, `variog.qn`.

Examples

```
# in S+ with S+SpatialStats:
if (!is.R()) {
vg1 <- variogram(coal~loc(x,y), data=coal.ash)
vg2 <- variogram.qn(coal.ash, lag.vect=vg1$distance,
          tol.lag = 0, minpairs = 6, printit = FALSE)
vg3 <- variogram(coal~loc(x,y), data=coal.ash,
          method="robust")

plot(vg1)
points(vg2$distance,vg2$gamma,col=3)
points(vg3$distance,vg3$gamma,pch=2,col=4)
}
```