

# The covRobust Package

February 16, 2008

**Title** Robust Covariance Estimation via Nearest Neighbor Cleaning

**Version** 1.0

**Author** Naisyin Wang <nwang@stat.tamu.edu> and Adrian Raftery <raftery@stat.washington.edu>  
with contributions from Chris Fraley <fraley@stat.washington.edu>

**Description** The `cov.nnve()` function for robust covariance estimation by the nearest neighbor variance estimation (NNVE) method of Wang and Raftery (2002, JASA)

**Maintainer** Naisyin Wang <nwang@stat.tamu.edu>

**License** GPL version 2 or newer

## R topics documented:

`cov.nnve` . . . . . 1

**Index** . . . . . 3

---

`cov.nnve`                      *Robust Covariance Estimation via Nearest Neighbor Cleaning*

---

## Description

The `cov.nnve` function for robust covariance estimation by the nearest neighbor variance estimation (NNVE) method of Wang and Raftery (2002, JASA).

## Usage

```
cov.nnve(datamat, k = 12, pnoise = 0.05, emconv = 0.001, bound = 1.5,  
         extension = TRUE, devsm = 0.01)
```

**Arguments**

<code>datamat</code>	matrix in which each row represents an observation or point and each column represents a variable
<code>k</code>	desired number of nearest neighbors (default is 12)
<code>pnoise</code>	percent of added noise
<code>emconv</code>	convergence tolerance for EM
<code>bound</code>	value used to identify surges in variance caused by outliers wrongly included as signal points ( <code>bound = 1.5</code> means a 50 percent increase)
<code>extension</code>	whether or not to continue after reaching the last chi-square distance. The default is to continue, which is indicated by setting <code>extension = TRUE</code> .
<code>devsm</code>	when <code>extension = TRUE</code> , the algorithm stops if the relative difference in variance is less than <code>devsm</code> . (default is 0.01)

**Value**

A list with the following components:

<code>cov</code>	covariance matrix
<code>mu</code>	mean vector
<code>postprob</code>	posterior probability
<code>classification</code>	classification (0=noise otherwise 1) obtained by rounding <code>postprob</code>
<code>inncc</code>	list of initial nearest neighbor cleaning results (components are the covariance, mean, posterior probability and classification)

**Note**

terms of use: GPL version 2 or newer.

**References**

Wang and Raftery (2002), Nearest neighbor variance estimation (NNVE): Robust covariance estimation via nearest neighbor cleaning (with discussion), *Journal of the American Statistical Association* 97:994-1019

see also University of Washington Statistics Technical Report 368 (2000) <http://www.stat.washington.edu/www/research/reports>

**Examples**

```
data(iris)
cov.nnve(iris[-5])
```

# Index

\*Topic **multivariate**

cov.nnve, 1

\*Topic **robust**

cov.nnve, 1

cov.nnve, 1