Amap Package

Antoine Lucas

November 17, 2005

Contents

1 Overview

`Amap` package includes standard hierarchical clustering and k-means. We optimize implementation (with a parallelized hierarchical clustering) and allow the possibility of using different distances like Eulidean or Spearman (rank-based metric).

We implement a principal component analysis (with robusts methods).

2 Usage

2.1 Clustering

The standard way of building a hierarchical clustering:

```
> library(amap)

Loading required package: Biobase

Welcome to Bioconductor

        Vignettes contain introductory material.
        To view, simply type 'openVignette()' or start with 'help(Biobase)'.
        For details on reading vignettes, see the openVignette help page.

> data(USArrests)
> h = hcluster(USArrests)
> plot(h)
```

Or for the "heatmap":

```
> heatmap(as.matrix(USArrests), hclustfun = hcluster, distfun = function(u) {
+      u
+ })
```

On a multiprocessor computer:

```
> h = hclusterpar(USArrests, nbproc = 4)
```

The K-means clustering:

```
> Kmeans(USArrests, centers = 3, method = "correlation")
```

2.2 Robust tools

A robust variance computation:

```
> data(lubisch)
> lubisch <- lubisch[, -c(1, 8)]
> varrob(scale(lubisch), h = 1)
```

A robust principal component analysis:

```
> p <- acpgen(lubisch, h1 = 1, h2 = 1/sqrt(2))
> plot(p)
```

Another robust pca:

```
> p <- acprob(lubisch, h = 4)
> plot(p)
```

3 See Also

Theses examples can be tested with command demo(amap).

All functions has got man pages, try help.start().

Robust tools has been published: [2] and [1].

References

[1] H. Caussinus, M. Fekri, S. Hakam, and A. Ruiz-Gazen. A monitoring display of multivariate outliers. *Computational Statistics and Data Analysis*, 44:237–252, October 2003.

[2] H. Caussinus, S. Hakam, and A. Ruiz-Gazen. Projections révélatrices contrôlées. recherche d'individus atypiques. *Revue de Statistique Appliquée*, 50(4), 2002.