Package 'ipr'

July 22, 2025

Type Package

Title Iterative Proportional Repartition Algorithm

Version 0.1.0

Author Dr. Jean-Benoit Rossel, Prof. Valentin Rousson, Dr. Yves Eggli

Maintainer Jean-Benoit Rossel <jean-benoit.rossel@unisante.ch>

Description Let us consider a sample of patients who can suffer from several diseases simultaneously, in a given set of diseases. The goal of the implemented algorithm is to estimate the individual average cost of each disease, starting from the global health costs available for each patient.

License GPL (>= 2)

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2020-03-20 14:10:02 UTC

Contents

---

ipr	*Iterative Proportional Repartition (IPR) algorithm*

Description

Estimating the health cost repartition among diseases in the presence of multimorbidity, i.e. when some patients have multiple diseases. Using the Iterative Proportional Repartition algorithm (see reference below), the goal is to estimate the average cost for each disease, starting from the global health costs available for each patient.

Usage

```
ipr(X, y, print.it=FALSE, start=rep(1,dim(X)[2]), cutup=Inf, cutlow=cutup,
epsrel=0.001, epsabs=0.1, maxiter=1000, det=FALSE)
```

Arguments

X	Matrix with $x_{ij} = 1$ if patient i suffers from disease j and $x_{ij} = 0$ otherwise. Each row thus refers to one patient and each column to one disease. The number of columns of X corresponds to the number of diseases considered.
y	Vector where y_i is the global health cost of patient i. The length of y must be equal to the number of rows of X.
print.it	Logical. If TRUE, the number of the current iteration and the current estimates are printed.
start	Vector of initial estimates of the average cost for each disease to start IPR algorithm. Default is an initial average cost of 1 for all diseases. The length of start must be equal to the number of columns of X.
cutup, cutlow	Options which can be used to get a robust version of IPR. If the current allocated cost of disease j for patient i is more than cutup times more expansive (or less then cutlow times less expansive) than the current average cost estimate of that disease j, then this outlying allocated cost is not taken into account in the next iteration to compute the average cost of disease j. By default, cutup and cutlow are set to Inf.
epsrel	Stopping criterion such that the IPR algorithm stops if for all diseases, the current estimated average cost differs by less than 100*epsrel percent from what it was at the previous iteration. The default value is 0.001. Should be set to 0 to ignore that criterion.
epsabs	Stopping criterion such that the IPR algorithm stops if for all diseases, the current estimated average cost differs (in absolute value) by less than epsabs percent from what it was at the previous iteration. The default value is 0.1. Should be set to 0 to ignore that criterion.
maxiter	Maximal number of iterations of IPR algorithm. The default value is 1000.
det	Logical. If TRUE, the allocated costs of each disease for each patient are given, by returning a matrix Y where y_{ij} is the estimated cost of disease j for patient i.

Details

Let us consider n patients and p diseases. We are given a matrix X such that $x_{ij} = 1$ if the patient i suffers from disease j and $x_{ij} = 0$ otherwise. We are also given a vector y, where y_i is the global health cost of patient i. In order to estimate the average cost of each disease, the IPR algorithm works as follows:

1. Start with some initial estimates mu_j, e.g. mu_j=1 for all $j = 1, \ldots, p$. Those initial estimates are stored in the vector start.

2. Allocate the cost y_i among the diseases diagnosed for patient i, proportionally to the current estimates mu_j.

3. Update the current estimate of mu_j by averaging the specific costs obtained in step 2 for the disease j over the patients having that disease.

4. Repeat steps 2 and 3 until a stopping criterion, based on relative or absolute distance between two consecutive iterations. The stopping criterion can be defined with epsabs or epsrel.

By construction, the IPR algorithm satisfies two properties. First, it allows to obtain positive estimates for each average disease cost. Secondly, it allows to retrieve the total health costs. In other words, the sum of the estimates mu_j multiplied by the number of patients suffering from j is equal to the sum of the costs y_i.

The estimate of total cost tau_j spent for disease j as well as the estimated proportion pi_j of the total costs which is allocated to disease j are also returned by our function.

Mathematically, tau_j is the sum over $i = 1$ to $i = n$ of $X_{ij} * mu_j$, while pi_j is defined by tau_j divided by the sum of all tau_k.

Value

coef	A vector with the estimated average cost of each disease.
total	A vector with the estimated total cost spent for each disease.
proportions	A vector with the estimated proportion of total cost spent for each disease.
niter	The number of iterations of IPR algorithm until the stopping criterion is achieved.
esprel	The stopping criterion based on a relative distance between two consecutive iterations which has been used.
epsabs	The stopping criterion based on an absolute distance between two consecutive iterations which has been used.
detail	A matrix with the allocated costs of each disease for each patient, if det is set to TRUE.

Author(s)

Dr. Jean-Benoit Rossel (<jean-benoit.rossel@unisante.ch>), Prof. Valentin Rousson and Dr. Yves Eggli.

References

Rousson, V., Rossel, J.-B. & Eggli, Y. (2019). Estimating Health Cost Repartition Among Diseases in the Presence of Multimorbidity. *Health Services Research and Managerial Epidemiology, 6.*

Rossel, J.-B., Rousson, V. & Eggli, Y. A comparison of statistical methods for allocating disease costs in the presence of interactions. In preparation.

Examples

```
# Here is a first example with 10 patients and 4 diseases:
X <- matrix(c(1,0,0,0,
0,1,1,0,
0,1,0,1,
1,0,0,1,
1,1,1,0,
```

```
0,0,1,1,
0,1,0,0,
1,1,0,0,
0,1,1,1,
0,0,0,1),ncol=4,byrow=TRUE)

y <- c(500,200,100,400,1000,500,100,300,800,2000)

# If we would use a linear model without intercept to estimate the average
# disease costs, we would obtain a negative value for disease 2.
lm(y~X-1)

# The IPR algorithm provides only positive estimates
ipr(X,y)


# Here is a second example:
X <- matrix(c(1,0,0,1,1,1),nrow=3,byrow=TRUE)
y <- c(5000,500,6600)

# We have three patients. The first one has only disease 1 with a cost of 5000.
# The second one has only disease 2 with a cost of 500 (i.e. ten times less
# expansive than disease 1). The third patient has both diseases with
# a cost of 6600 (i.e. 5000 + 500 + an extra cost of 1100).

# Using a linear model, one would allocate the extra cost equally between
# the three patients. The estimated average cost would thus be 5000+(1100/3)
# for disease 1 and 500+(1100/3) for disease 2.
lm(y~X-1)

# Using IPR algorithm, one allocates the extra cost taking into account that
# disease 1 is ten times more expansive than disease 2 when occuring alone.
# One thus gets an estimated average cost of 5500 for disease 1 and
# of 550 for disease 2.
ipr(X,y)
```

Index