

Package ‘p2distance’

September 15, 2026

Type Package

Title Welfare's Synthetic Indicator

Version 1.0.2

Date 2026-08-24

Description The welfare's synthetic indicator provides an ideal tool for measuring multi-dimensional concepts such as welfare, development, living standards, etc. It enables information from the various indicators to be aggregated into a single synthetic measure. The method was proposed by Pena (1977, ISBN:9788426001788) and further developed by Zarzosa and Somarriba (2012) <[doi:10.1007/s11205-012-0005-0](https://doi.org/10.1007/s11205-012-0005-0)>.

License GPL

LazyLoad yes

URL <https://github.com/ajpelu/p2distance>,
<https://ajpelu.github.io/p2distance/>

BugReports <https://github.com/ajpelu/p2distance/issues>

Encoding UTF-8

RoxygenNote 8.0.0

Config/Needs/website rmarkdown

Suggests cranlogs, dplyr, ggplot2, lubridate, tidyr

NeedsCompilation no

Author Antonio J. Pérez-Luque [aut, cre] (ORCID:
<<https://orcid.org/0000-0002-1747-0469>>),
Ricardo Moreno [ctb],
Ramón Pérez-Pérez [aut],
Francisco J. Bonet-García [ctb]

Maintainer Antonio J. Pérez-Luque <ajpelu@gmail.com>

Repository CRAN

Date/Publication 2026-09-15 11:50:02 UTC

Contents

loadCSVtoP2distance	2
makeReferenceVector	3
p2distance	4
welfare	6
Index	9

loadCSVtoP2distance	<i>Load a CSV file into a matrix object</i>
---------------------	---

Description

Reads a text file and converts it into the matrix object expected by `p2distance()`.

Usage

```
loadCSVtoP2distance(  
  path,  
  header = TRUE,  
  sep = "\t",  
  dec = ".",  
  quote = "\"",  
  na.strings = "NA",  
  fileEncoding = "",  
  encoding = "unknown"  
)
```

Arguments

path	The path of the file from which the data are to be read. Each row of the table appears as one line of the file.
header	A logical value indicating whether the file contains the names of the variables as its first line.
sep	The field separator character.
dec	The character used in the file for decimal points.
quote	The set of quoting characters. To disable quoting altogether, use <code>quote = ""</code> .
na.strings	A character vector of strings to be interpreted as NA values.
fileEncoding	Character string: if non-empty, declares the encoding used in the file so the character data can be re-encoded. See the "Encoding" section of <code>?file</code> .
encoding	Encoding to be assumed for input strings.

Details

The first column of the file is used to set the row names of the resulting matrix (typically the name of each spatial entity), and is then removed from the data itself. Internally this uses `utils::read.table()` to read the file.

Value

A matrix containing the data from the CSV file, ready to be passed to [p2distance\(\)](#).

See Also

[p2distance\(\)](#)

makeReferenceVector	<i>Make a reference vector</i>
---------------------	--------------------------------

Description

A reference vector must be defined for each partial indicator so as to compare different spatial entities. This vector is used by [p2distance\(\)](#) (as the base reference) to calculate the distance of each spatial entity from this reference base. It is common to use the minimum value as the base reference (see references).

Usage

```
makeReferenceVector(X, reference_vector_function = min)
```

Arguments

X	A data frame or matrix with n rows (entities) and m columns (variables).
reference_vector_function	Function used to build the reference vector. Common choices are min, max, mean, median, etc. See apply() for details. min is the default.

Value

A vector with the reference value (base reference) for each variable. Its length equals the number of variables.

References

- Peña, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.
- Peña, J. B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, 27(2), 299-324.
- Somarriba, N. (2008). *Aproximación a la medición de la calidad de vida en la Unión Europea*. Doctoral Thesis. University of Valladolid.
- Zarzosa, P. (1992). *Aproximación a la medición del bienestar social, estudio de la idoneidad del indicador sintético Distancia P2*. Doctoral Thesis. University of Valladolid.

See Also

[p2distance\(\)](#)

Examples

```
## Create a data frame of 3 variables (indicators) for 6 entities (rows)
dat <- data.frame(
  x1 = c(10, 12, 13, 14, 12, 11),
  x2 = c(40, 51, 61, 68, 34, 44),
  x3 = c(0.43, 0.41, 0.39, 0.55, 0.60, 0.38)
)

## Calculate the reference vector using the minimum value of each variable
makeReferenceVector(dat, reference_vector_function = min)
```

p2distance	<i>Welfare's synthetic indicator (P2 distance)</i>
------------	--

Description

Calculates the P_2 distance synthetic indicator for a set of variables. This is the main function of the package.

Usage

```
p2distance(
  matriz,
  reference_vector = NULL,
  reference_vector_function = min,
  iterations = 20,
  umbral = 1e-04
)
```

Arguments

matriz	A matrix with spatial entities in rows and variables in columns.
reference_vector	Optional. A reference vector defined for each partial indicator, used to compare different spatial entities.
reference_vector_function	Optional. Function used to build the reference vector when <code>reference_vector</code> is not supplied. <code>min</code> is the default; other common choices are <code>max</code> , <code>mean</code> , <code>median</code> , etc. See makeReferenceVector() for details.
iterations	Maximum number of iterations for the computational process until convergence is reached.
umbral	The algorithm stops when the difference between two consecutive iterations is lower than this threshold.

Details

The P_2 distance, also called DP2, is used to measure welfare in quality-of-life applications, to build environmental quality indexes, and more generally to aggregate multiple partial indicators (variables) into a single measure that allows spatial entities to be compared. For a spatial entity r , the P_2 distance is defined as:

$$DP_2 = \sum_{i=1}^n \left\{ \left(\frac{d_i}{\sigma_i} \right) (1 - R_{i,i-1,i-2,\dots,1}^2) \right\}$$

with $R_1^2 = 0$, where $d_i = |x_{ri} - x_{*i}|$, with the reference base $X_* = (x_{*1}, x_{*2}, \dots, x_{*n})$, and:

- n is the number of variables
- x_{ri} is the value of variable i for spatial entity r
- σ_i is the standard deviation of variable i
- $R_{i,i-1,\dots,1}^2$ is the coefficient of determination of the regression of X_i on $X_{i-1}, X_{i-2}, \dots, X_1$ already included

The numerical value of the DP2 index has no meaning by itself, but it is useful for comparing the state of different spatial entities in terms of welfare, environmental conditions, etc.

Value

A list with the following elements:

- `discrimination.coefficient`: Vector of discrimination coefficients (DC) for each variable (Ivanovic, 1974). DC ranges between 0 and 2: a variable with the same value for every spatial entity has $DC = 0$ (no discriminant power), while a variable with a single non-zero value has $DC = 2$ (full discriminant power). See Zarzosa (1996) and Zarzosa & Somarriba (2012).
- `p2distance`: Vector with the final P_2 distance value for each spatial entity.
- `p2distances`: Matrix with the P_2 distance values resulting from each iteration.
- `diff_p2distances`: Matrix with the differences between two consecutive P_2 distances.
- `iteration`: Number of iterations performed.
- `umbral`: Threshold used to stop the iterations.
- `variables_sort`: Variable names ordered by entrance order in the last iteration.
- `correction_factors`: Correction factor for each variable.
- `cor.coeff`: Correlation coefficient of each variable with the calculated P_2 distance.
- `partial.Indicators`: For each spatial entity, the difference between the reference vector and the value of each variable, divided by the standard deviation. The sum of all partial indicators for a spatial entity is the Frechet Distance (DF), the maximum value the P_2 distance can reach.

References

Ivanovic, B. (1974). Comment établir une liste des indicateurs de developpement. *Revue de Statistique Appliquée*, 22(2), 37-50.

Montero, J. M., Chasco, C., & Larraz, B. (2010). Building an environmental quality index for a big city: a spatial interpolation approach combined with a distance indicator. *Journal of Geographical Systems*, 12, 435-459.

Peña, J. B. (1977). *Problemas de la medición del bienestar y conceptos afines (una aplicación al caso Español)*. Madrid: INE.

Peña, J. B. (2009). La medición del bienestar social: una revisión crítica. *Estudios de Economía Aplicada*, 27(2), 299-324.

Zarzosa, P. (1996). *Aproximación a la medición del Bienestar social*. Valladolid: Universidad de Valladolid.

Zarzosa, P., & Somarriba, N. (2012). An assessment of social welfare in Spain: Territorial analysis using a synthetic welfare indicator. *Social Indicators Research*. doi:10.1007/s1120501200050

See Also

[makeReferenceVector\(\)](#), [loadCSVtoP2distance\(\)](#)

Examples

```
## Calculate a welfare indicator for 27 European countries
data(welfare)
welfare <- as.matrix(welfare)

ind <- p2distance(welfare, reference_vector_function = min, iterations = 20)

## Examine the results
ind$p2distance
ind$iteration
ind$variables_sort
ind$correction_factors
ind$cor.coeff
ind$discrimination.coefficient

## Plot of the P2 distance indicator for European countries
barplot(ind$p2distance, beside = TRUE, col = "white", space = .3,
        ylab = "P2 distance", ylim = c(0, 20),
        names.arg = rownames(ind$p2distance), las = 3, cex.names = 0.8)
```

welfare

Welfare indicators of European countries

Description

Quality of life indicators (*welfare indicators*) for 27 countries of the European Union during the 2002-2007 period. For each variable (i.e. indicator), both its definition and the year it refers to are given. The data come from EurLIFE, an interactive database on quality of life in Europe (see Source).

Usage

```
data(welfare)
```

Format

A data frame with 27 rows and 20 columns:

- `happiness`: Mean value on a scale from 1 ("very unhappy") to 10 ("very happy"). Year: 2007.
- `life.satis`: Life satisfaction. Share of individuals aged 15 and over who are very or fairly satisfied with the life they lead. Year: 2006.
- `judicial`: Trust in the judicial system. Percentage of people aged 15 and over who tend to trust justice or the legal system. Year: 2005.
- `night`: Unsafe to walk around at night. Percentage of individuals aged 18 and over who think it is very or rather unsafe to walk around the area they live in at night. Year: 2003.
- `social`: Satisfaction with social life. Percentage of the population aged 15 and over who are very or fairly satisfied with their social life. Year: 2004.
- `people`: Trust in people. Mean value on a scale from 1 ("you can't be too careful in dealing with people") to 10 ("most people can be trusted"). Year: 2007.
- `family`: Satisfaction with family life. Mean value on a scale from 1 ("very dissatisfied") to 10 ("very satisfied"). Year: 2007.
- `health`: Satisfaction with the national health care system. Percentage of individuals who are very or fairly satisfied with their national health system. Year: 2002.
- `life.65`: Life expectancy at age 65. Average number of further years a person aged 65 would live if age-specific mortality rates remained constant. Year: 2003.
- `life.0`: Life expectancy at birth. Average number of years a person would live if age-specific mortality rates remained constant. Year: 2005.
- `inequality`: Inequality of income distribution. Ratio of the total income received by the 20% of the population with the highest income to that received by the 20% with the lowest income. Year: 2005.
- `hobbies`: Too little time for hobbies and interests. Percentage of people aged 18 and over who have too little time for hobbies and interests. Year: 2007.
- `education`: Satisfaction with education. Mean value on a scale from 1 ("very dissatisfied") to 10 ("very satisfied"). Year: 2007.
- `standard`: Satisfaction with standard of living. Mean value on a scale from 1 ("very dissatisfied") to 10 ("very satisfied"). Year: 2007.
- `dist.school`: Distance to the nearest primary school. Proportion of people aged 15 and over who live within walking distance or within 20 minutes of the nearest primary school. Years: 1999, 2005.
- `area`: Satisfaction with the area you live in. Share of individuals aged 15 and over who are very or fairly satisfied with the area they live in (five-item scale). Year: 2004.
- `home`: Satisfaction with the home. Share of individuals aged 15 and over who are very or fairly satisfied with their home (four-item scale). Year: 2004.
- `stress`: Find work stressful. Share of people who "strongly agree" or "agree" that their work is too demanding and stressful. Year: 2007.

- employment: Employment rate. Employed persons aged 15-64 as a percentage of the total population in that age group. Year: 2005.
- job: Job satisfaction. Percentage of employed people who are very or fairly satisfied with their job (four-item scale). Year: 2005.

Details

These quality of life indicators come from different sources (Standard Eurobarometer Surveys, European Quality of Life Surveys, Eurostat, European Foundation for the Improvement of Living and Working Conditions) collated by EurLIFE.

Source

EurLIFE database. Interactive database on quality of life in Europe. European Foundation for the Improvement of Living and Working Conditions (Eurofound). Accessed 20 Feb 2012.

References

Somarriba, N., & Peña, B. (2009). Synthetic Indicators of Quality of Life in Europe. *Social Indicators Research*, 94, 115-133.

Examples

```
data(welfare)
```

Index

* **datasets**

welfare, [6](#)

apply(), [3](#)

loadCSVtoP2distance, [2](#)

loadCSVtoP2distance(), [6](#)

makeReferenceVector, [3](#)

makeReferenceVector(), [4](#), [6](#)

p2distance, [4](#)

p2distance(), [2](#), [3](#)

utils::read.table(), [2](#)

welfare, [6](#)