

# Package: earthmovdist (via r-universe)

July 7, 2024

**Version** 0.1.3

**Date** 2017-10-08

**Author** Rainer M Krug <Rainer@krugs.de> and Dirk Eddelbuettel  
<edd@debian.org>

**Maintainer** Dirk Eddelbuettel <edd@debian.org>

**Title** Wrapper to the 'Emd-L1' Library by Haibin Ling and Kazunori  
Okada

**Description** The 'Emd-L1' distance metric from the paper by H. Ling and  
K. Okada, An Efficient Earth Mover's Distance Algorithm for  
Robust Histogram Comparison, IEEE Transaction on Pattern  
Analysis and Machine Intelligence (PAMI), 29(5):840-853, 2007,  
is provided for use by R. It is based on code downloaded from  
the URL <[http://www.ist.temple.edu/~hbling/code/EmdL1\\_v3.zip](http://www.ist.temple.edu/~hbling/code/EmdL1_v3.zip)>.

**Depends** Rcpp (>= 0.8.0)

**LinkingTo** Rcpp

**License** file LICENSE

**URL** <http://earthmovdist.r-forge.r-project.org/>

**Repository** <https://eddelbuettel.r-universe.dev>

**RemoteUrl** <https://github.com/eddelbuettel/earthmovdist>

**RemoteRef** HEAD

**RemoteSha** 8950bdc869ed922f411c6607a1b2fdd4d8196827

## Contents

emdL1 . . . . .	2
<b>Index</b>	<b>4</b>

emdL1

*Earth Mover's Distance between histograms***Description**

The emdL1 provides a wrapper to the emd-L1 distance norm by Haibin Ling and Kazunori Okada.

**Usage**

```
emdL1(x, y, dims=NULL, verbose=FALSE)
```

**Arguments**

x	vector, matrix or 3D array containing signature 1
y	vector, matrix or 3D array containing signature 2
dims	optional vector of dimensions of input data, see Details
verbose	optional boolean flag indicating verbose operation

**Details**

This function provides an implementation of the Earth Movers Distance (EMD), using the L1 (taxi-cab) metric.

To quote wikipedia ([http://en.wikipedia.org/wiki/Earth\\_Mover's\\_Distance](http://en.wikipedia.org/wiki/Earth_Mover's_Distance)):

'The earth mover's distance (EMD) is a mathematical measure of the distance between two distributions over some region D. Informally, if the distributions are interpreted as two different ways of piling up a certain amount of dirt over the region D, the EMD is the minimum cost of turning one pile into the other; where the cost is assumed to be amount of dirt moved times the distance by which its is moved'

(For a detailed discussion of the EMD, see [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/RUBNER/emd.htm](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/RUBNER/emd.htm))

The distance 'the dirt' has to be moved is in the implementation of emd-L1, the L1 distance: if we have two points (x1, y1) and (x2, y2), the L1 distance (or taxicab metric) is  $\text{abs}(x2-x1)+\text{abs}(y2-y1)$ .

The above definition only covers the case, where the 'amount of dirt' in both distributions is the same. In this case, the EMD is identical to the 1st Mallows distance or 1st Wasserstein distance.

For details on how the emd-L1 distance is computed in this algorithm, see

```
@ARTICLE{
  author = {Ling, Haibin and Okada, K.},
  title = {An efficient earth mover's distance algorithm for robust histogram
    comparison},
  journal = {IEEE Transactions on Pattern Analysis and Machine Intelligence},
  year = {2007},
  volume = {29},
  pages = {840-53},
```

```
number = {5},  
month = {05},  
doi = {10.1109/TPAMI.2007.1058},  
publication-type = {J},  
publisher = {IEEE Comput. Soc},  
type = {Journal Paper}  
}
```

The function has to be called by providing two signatures ( $x$  and  $y$ ). these can be vectors (1D signatures), matrix (2D signatures) or 3D arrays (3D signatures). The dimensions of  $x$  and  $y$  have to be the same, and  $x$  and  $y$  should not contain any missing or infinite values.

If the vector `dims` is supplied,  $x$  and  $y$  have to be vectors, and `dims` will be used to interpret those as vectors (`length(dims)==1`), matrix (`length(dims)==2`) or 3D array (`length(dims)==3`).

If the vector `dims` is not supplied and `dim(x)==dim(y)`, `dims` is set internally to `dim(x)`.

### Value

A real number with the EMD-L1 distance metric between the two histograms

### Author(s)

Dirk Eddelbuettel and Rainer M Krug (Rainer@krugs.de) for the R package; Haibin Ling and Kazunori Okada for emd-L1.

### Examples

```
set.seed(42)  
x <- rnorm(100)  
y <- x + rnorm(100)/10  
emdL1(x, y, verbose=TRUE)
```

# Index

emdL1, [2](#)