Read me

The folders of data files contain most of the data sets used in *Analysis of multivariate and high-dimensional data* (AMHD). Some data sets are not available or are only available on request, and I will comment on these below.

The names of the folders are those used in the Data Index of AMHD. Each folder refers to one data set which is typically given as a .csv file. A folder will contain extra information relating to the data. For some data sets the names of the variables are given in the csv file. In some cases the data are given in a different format, eg the sound tracks data are given as .wav files, and the PBMC flow cytometry data are given as .fsc files. In each case I have supplied comments on how to read these data into Matlab.

Data sets in the AMHD Data Index which are not included:

Breast tumour: only the version with d=4751 and n=78 is included. If the reader requires the file with d=24,48, he/she may contact the author directly.

Car: available from within Matlab by using the commands
```
    load carbig;
    xin = [Displacement Horsepower Weight Acceleration MPG];
    nans = sum(isnan(xin),2) > 0;
    inmat = (xin(~nans,1:5))';
```
the last 3 lines of Matlab code find missing values, delete the rows with missing value and return a matrix of size 5x392.

Cereal: available from within Matlab by using the commands
```
    load cereal;
    xin = [Calories Protein Fat Sodium Fiber ...
    Carbo Sugars Shelf Potass Vitamins Weight];
    inmat = (X(strmatch('K',Mfg),:))';
```
the last 3 lines of Matlab code define the 77x11 input data, and inmat is the 11x23 subset which corresponds to the single manufacturer Kellogg's.

Exam grades: available from within Matlab, and results in a 5x120 matrix:
```
    load examgrades;
    inmat = grades';
```

HRCT emphysema: not publicly available due to confidentiality issues of the collectors of these data.

(data bank of) kidneys: please contact the author directly for these data.

Ovarian cancer proteomics: please contact the author directly for these data.

Pitprops: not publicly available. See Jeffers (1967) in the bibliography of AMHD for more detail.

South Australian grapevine: please contact the owner of the data, Chris Davies, directly. For contact details see the footnote of Example 6.7.

Inge Koch
November 2013