

DATA EXTRACTION RAW DOW-JONES.

LAURENT CALLOT

This file extracts the Dow Jones series, un-transformed and un-censored for the master thesis of Maxime Apeldoorn. We have 1474 daily observations for 465 series which correspond to the unique entries of the covariance matrix of the 30 stocks of the Dow Jones.

The data is a subset of the S&P 500 realized covariance matrices estimated in Barndorff-Nielsen, Hansen, Lunde, and Shephard (Econometrica 2008), the original data is kindly provided by Asger Lunde.

Let y_t be a row of the data, and Σ_t the realized covariance matrix at time t , and $\sigma_{t,i,j}$ the entry at the i^{th} row and j^{th} column of Σ_t . Note that by symmetry $\sigma_{t,i,j} = \sigma_{t,j,i}$. We have $n = 30$ stocks so that Σ_t is $[n \times n]$ and y_t has length $\frac{n(n+1)}{2}$. y_t is constructed as $y_t = vech(\Sigma_t)$, meaning that:

$$y_t = [\sigma_{t,1,1}, \sigma_{t,1,2}, \sigma_{t,2,2}, \sigma_{t,1,3}, \dots, \sigma_{t,n-1,n}, \sigma_{t,n,n}].$$

The data comprises 3 files:

- (1) `dj_crk_raw.txt` is a plain text files containing 1474 rows of values separated by spaces. Each row contains a date in the format YYYY-MM-DD followed by a y_t .
- (2) `dj_crk_dates.txt` a plain text file containing a single column with the dates.
- (3) `dj_crk_names.txt` contains the ticker for each of the 30 stocks in the order in which they appear in Σ_t as well as information on the industry categories to which they belong.

```
##          spec
## Data    Stocks Censoring Transformation
##   [1,] "dj"    "none"    "none"
## .....
## Extracting data for stock set: dj . Contains 30
## stocks.
## Stock Names and Index:
##   AA 98
##   MO 27
##   AXP 225
##   T 259
##   BAC 228
##   BA 57
##   CAT 51
##   CVX 31
##   CSCD 302
##   KO 9
##   DD 13
##   XOM 16
##   GE 20
```

```
## HPQ 118
## HD 262
## INTC 227
## IBM 23
## JNJ 72
## JPM 186
## KFT 417
## MCD 168
## MRK 77
## MSFT 2
## PFE 69
## PG 48
## UTX 47
## VZ 257
## WMT 214
## DIS 110
## MMM 76
##
##
## .....
## Data Saving:
##
## Total number of specifications: 1
##
## Stock Index: dj, Censoring: none, Transformation: none.
## File names: CRK.dj.none.none
##
## CRK.har.dj.none.none
##
##
## And we're done
##
## Time used to generate the files, in seconds:
## 13.49
```