Intro to `DatABEL`

Yurii S. Aulchenko

April 4, 2010

Contents

```
> library(DatABEL)
> make_random_matrix <- function(range_dim1 = c(2, 10), range_dim2 = c(2,
+     10), range_data = c(-10, 10), type = "double") {
+     dim1 <- round(runif(1, range_dim1[1], range_dim1[2]))
+     dim2 <- round(runif(1, range_dim2[1], range_dim2[2]))
+     data <- runif(dim1 * dim2, range_data[1], range_data[2])
+     data <- as(data, type)
+     data <- matrix(data, nrow = dim1, ncol = dim2)
+     namesCol <- paste("col", c(1:dim2), sep = "_")
+     namesRow <- paste("row", c(1:dim1), sep = "_")
+     dimnames(data) <- list(namesRow, namesCol)
+     return(data)
+ }
> testmatr <- make_random_matrix()
> testmatr

          col_1       col_2          col_3
row_1 -7.460411 -9.2190394   0.001238710
row_2  9.839267  5.4834475  -6.926002498
row_3  9.654159 -4.7130789  -0.399104827
row_4 -2.267890  9.9433627  -9.739416051
row_5  3.855142 -8.1034728  -3.419599105
row_6 -2.878103 -6.2862648   8.396422463
row_7 -2.262509  0.5046935   9.494556366

> test_fv <- as(testmatr, "databel_base_R")

Initializing empty file './tmp676577', type 6.
type = 6(DOUBLE)
nelements = 21
numObservations = 7
numVariables = 3;
bytesPerRecord = 8;
```

1

```
bitsPerRecord = 64;
File './tmp676577' initialized.
[1] "./tmp676577"
Opening FileVector './tmp676577'.
You appear to work on 32-bit system. Large files are not supported.
Filevector ./tmp676577 opened.
Opening FileVector './tmp676577'.
You appear to work on 32-bit system. Large files are not supported.
Filevector ./tmp676577 opened.
coersion from 'matrix' to 'databel_base_R' of type DOUBLE ; object connected to file ./tmp67
```

```
> test_fv

backingfilename = ./tmp676577
cachesizeMb = 1
number of columns (variables) =  3
number of rows (observations) =  7
Upper-left 3 columns and  5 rows:
          [,1]      [,2]         [,3]
[1,] -7.460411 -9.219039  0.001238710
[2,]  9.839267  5.483448 -6.926002502
[3,]  9.654159 -4.713079 -0.399104834
[4,] -2.267889  9.943362 -9.739416122
[5,]  3.855142 -8.103473 -3.419599056
```

```
> as(test_fv, "matrix")

           [,1]       [,2]         [,3]
[1,] -7.460411 -9.2190390  0.001238710
[2,]  9.839267  5.4834476 -6.926002502
[3,]  9.654159 -4.7130790 -0.399104834
[4,] -2.267889  9.9433622 -9.739416122
[5,]  3.855142 -8.1034727 -3.419599056
[6,] -2.878103 -6.2862649  8.396422386
[7,] -2.262509  0.5046936  9.494556427
```

```
> abs(testmatr - as(test_fv, "matrix")) < 1e-06

      col_1 col_2 col_3
row_1  TRUE  TRUE  TRUE
row_2  TRUE  TRUE  TRUE
row_3  TRUE  TRUE  TRUE
row_4  TRUE  TRUE  TRUE
row_5  TRUE  TRUE  TRUE
row_6  TRUE  TRUE  TRUE
row_7  TRUE  TRUE  TRUE
```

```
> write.table(testmatr, file = "test_matrix_dimnames.dat", row.names = TRUE,
+     col.names = TRUE, quote = FALSE)
> text2filevector(infile = "test_matrix_dimnames.dat", outfile = "test_matrix_dimnames",
+     R_matrix = TRUE)

Options in effect:
        --infile    = test_matrix_dimnames.dat
        --outfile   = test_matrix_dimnames
        --skiprows  = 1
        --skipcols  = 1
        --cnrow     = ON, using line 1 of 'test_matrix_dimnames.dat'
        --rncol     = ON, using column 1 of 'test_matrix_dimnames.dat'
        --transpose = OFF
        --Rmatrix   = ON
Creating file with numRows = 7
Creating file with numColumns = 3
Initializing empty file 'test_matrix_dimnames_fvtmp', type 6.
type = 6(DOUBLE)
nelements = 21
numObservations = 3
numVariables = 7;
bytesPerRecord = 8;
bitsPerRecord = 64;
File 'test_matrix_dimnames_fvtmp' initialized.
Opening FileVector 'test_matrix_dimnames_fvtmp'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames_fvtmp opened.
Closing FileVector
Opening FileVector 'test_matrix_dimnames_fvtmp'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames_fvtmp opened.
Initializing empty file 'test_matrix_dimnames', type 6.
type = 6(DOUBLE)
nelements = 21
numObservations = 7
numVariables = 3;
bytesPerRecord = 8;
bitsPerRecord = 64;
File 'test_matrix_dimnames' initialized.
Opening FileVector 'test_matrix_dimnames'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames opened.
Copying var/obs names...Closing FileVector
Closing FileVector
done
Copying data...3x7
```

```
data written
done
text2fvf finished.
Opening FileVector 'test_matrix_dimnames'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames opened.
unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames
cachesizeMb = 1
number of columns (variables) =  3
number of rows (observations) =  7
usedRowIndex: 1  2  3  4  5  ...
usedColIndex: 1  2  3
Upper-left 3 columns and  5 rows:
            [,1]      [,2]           [,3]
[1,] -7.460411 -9.219039  0.001238710
[2,]  9.839267  5.483448 -6.926002502
[3,]  9.654159 -4.713079 -0.399104834
[4,] -2.267889  9.943362 -9.739416122
[5,]  3.855142 -8.103473 -3.419599056

> x <- databel_filtered_R("test_matrix_dimnames")

Opening FileVector 'test_matrix_dimnames'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames opened.

> x

unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames
cachesizeMb = 1
number of columns (variables) =  3
number of rows (observations) =  7
usedRowIndex: 1  2  3  4  5  ...
usedColIndex: 1  2  3
Upper-left 3 columns and  5 rows:
            [,1]      [,2]           [,3]
[1,] -7.460411 -9.219039  0.001238710
[2,]  9.839267  5.483448 -6.926002502
[3,]  9.654159 -4.713079 -0.399104834
[4,] -2.267889  9.943362 -9.739416122
[5,]  3.855142 -8.103473 -3.419599056
```

```
> tmp <- as(x, "matrix")
> tmp

          col_1       col_2         col_3
row_1 -7.460411 -9.2190390  0.001238710
row_2  9.839267  5.4834476 -6.926002502
row_3  9.654159 -4.7130790 -0.399104834
row_4 -2.267889  9.9433622 -9.739416122
row_5  3.855142 -8.1034727 -3.419599056
row_6 -2.878103 -6.2862649  8.396422386
row_7 -2.262509  0.5046936  9.494556427

> abs(testmatr - tmp) < 1e-06

      col_1 col_2 col_3
row_1  TRUE  TRUE  TRUE
row_2  TRUE  TRUE  TRUE
row_3  TRUE  TRUE  TRUE
row_4  TRUE  TRUE  TRUE
row_5  TRUE  TRUE  TRUE
row_6  TRUE  TRUE  TRUE
row_7  TRUE  TRUE  TRUE

> text2filevector(infile = "test_matrix_dimnames.dat", outfile = "test_matrix_dimnames_T",
+     R_matrix = TRUE, transpose = TRUE)

Options in effect:
        --infile   = test_matrix_dimnames.dat
        --outfile  = test_matrix_dimnames_T
        --skiprows = 1
        --skipcols = 1
        --cnrow    = ON, using line 1 of 'test_matrix_dimnames.dat'
        --rncol    = ON, using column 1 of 'test_matrix_dimnames.dat'
        --transpose = ON
        --Rmatrix  = ON
Creating file with numRows = 7
Creating file with numColumns = 3
Initializing empty file 'test_matrix_dimnames_T', type 6.
type = 6(DOUBLE)
nelements = 21
numObservations = 3
numVariables = 7;
bytesPerRecord = 8;
bitsPerRecord = 64;
File 'test_matrix_dimnames_T' initialized.
Opening FileVector 'test_matrix_dimnames_T'.
You appear to work on 32-bit system. Large files are not supported.
```

```
Filevector test_matrix_dimnames_T opened.
Closing FileVector
text2fvf finished.
Opening FileVector 'test_matrix_dimnames_T'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames_T opened.
unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames_T
cachesizeMb = 1
number of columns (variables) =  7
number of rows (observations) =  3
usedRowIndex: 1  2  3
usedColIndex: 1  2  3  4  5  6  7
Upper-left 7 columns and  3 rows:
              [,1]       [,2]       [,3]       [,4]      [,5]       [,6]       [,7]
[1,] -7.460410595   9.839267   9.6541586  -2.267889   3.855142  -2.878103  -2.2625091
[2,] -9.219038963   5.483448  -4.7130790   9.943362  -8.103473  -6.286265   0.5046936
[3,]  0.001238710  -6.926003  -0.3991048  -9.739416  -3.419599   8.396422   9.4945564

> x <- databel_filtered_R("test_matrix_dimnames_T")

Opening FileVector 'test_matrix_dimnames_T'.
You appear to work on 32-bit system. Large files are not supported.
Filevector test_matrix_dimnames_T opened.

> t(testmatr)

            row_1      row_2       row_3      row_4      row_5      row_6
col_1 -7.460410786   9.839267   9.6541588  -2.267890   3.855142  -2.878103
col_2 -9.219039376   5.483447  -4.7130789   9.943363  -8.103473  -6.286265
col_3  0.001238710  -6.926002  -0.3991048  -9.739416  -3.419599   8.396422
          row_7
col_1 -2.2625092
col_2  0.5046935
col_3  9.4945564

> x

unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames_T
cachesizeMb = 1
number of columns (variables) =  7
number of rows (observations) =  3
```

```
usedRowIndex: 1  2  3
usedColIndex: 1  2  3  4  5  6  7
Upper-left 7 columns and  3 rows:
finalizing AbstractMatrix: 0x14f8800
Closing FileVector
finalizing AbstractMatrix: 0x10d9a00
Closing FileVector
finalizing AbstractMatrix: 0x111c600
Closing FileVector
             [,1]      [,2]        [,3]        [,4]       [,5]       [,6]       [,7]
[1,] -7.460410595  9.839267  9.6541586 -2.267889  3.855142 -2.878103 -2.2625091
[2,] -9.219038963  5.483448 -4.7130790  9.943362 -8.103473 -6.286265  0.5046936
[3,]  0.001238710 -6.926003 -0.3991048 -9.739416 -3.419599  8.396422  9.4945564

> tmp <- as(x, "matrix")
> tmp

            row_1      row_2      row_3      row_4      row_5      row_6
col_1 -7.460410595  9.839267  9.6541586 -2.267889  3.855142 -2.878103
col_2 -9.219038963  5.483448 -4.7130790  9.943362 -8.103473 -6.286265
col_3  0.001238710 -6.926003 -0.3991048 -9.739416 -3.419599  8.396422
           row_7
col_1 -2.2625091
col_2  0.5046936
col_3  9.4945564

> abs(t(testmatr) - tmp) < 1e-06

      row_1 row_2 row_3 row_4 row_5 row_6 row_7
col_1  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
col_2  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
col_3  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE

> unlink("*.fv?")
> unlink("test_matrix_*")
```