Intro to DatABEL

Yurii S. Aulchenko

March 7, 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    col_4     col_5
row_1 -8.510850 -7.371706 -0.5829301 6.774361  1.793244
row_2 -5.033636  9.213437  0.1687509 5.115450 -5.191560

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

[1] "./tmp269886"
WARNING: you appear to work on 32-bit system... large files not supported
WARNING: you appear to work on 32-bit system... large files not supported
finalizing AbstractMatrix: 0x16ac800
coersion from 'matrix' to 'databel_base_R' of type DOUBLE ; object connected to file ./tmp26

> test_fv

backingfilename = ./tmp269886
cachesizeMb = 1
number of columns (variables) =  5
number of rows (observations) =  2
```

```
Upper-left 5 columns and  2 rows:
          [,1]      [,2]       [,3]     [,4]      [,5]
[1,] -8.510850 -7.371706 -0.5829301 6.774361  1.793244
[2,] -5.033636  9.213437  0.1687509 5.115449 -5.191560
```

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

```
          [,1]      [,2]       [,3]     [,4]      [,5]
[1,] -8.510850 -7.371706 -0.5829301 6.774361  1.793244
[2,] -5.033636  9.213437  0.1687509 5.115449 -5.191560
```

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

```
      col_1 col_2 col_3 col_4 col_5
row_1  TRUE  TRUE  TRUE  TRUE  TRUE
row_2  TRUE  TRUE  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
WARNING: you appear to work on 32-bit system... large files not supported
WARNING: you appear to work on 32-bit system... large files not supported
WARNING: you appear to work on 32-bit system... large files not supported
WARNING: you appear to work on 32-bit system... large files not supported
unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames
cachesizeMb = 1
number of columns (variables) =  5
number of rows (observations) =  2
usedRowIndex: 1  2
usedColIndex: 1  2  3  4  5
Upper-left 5 columns and  2 rows:
          [,1]      [,2]       [,3]     [,4]      [,5]
[1,] -8.510850 -7.371706 -0.5829301 6.774361  1.793244
[2,] -5.033636  9.213437  0.1687509 5.115449 -5.191560
```

```
> x <- databel_filtered_R("test_matrix_dimnames")

WARNING: you appear to work on 32-bit system... large files not supported

> x

unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames
cachesizeMb = 1
number of columns (variables) =  5
number of rows (observations) =  2
usedRowIndex: 1  2
usedColIndex: 1  2  3  4  5
Upper-left 5 columns and  2 rows:
          [,1]      [,2]       [,3]     [,4]      [,5]
[1,] -8.510850 -7.371706 -0.5829301 6.774361  1.793244
[2,] -5.033636  9.213437  0.1687509 5.115449 -5.191560

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

          col_1     col_2      col_3    col_4     col_5
row_1 -8.510850 -7.371706 -0.5829301 6.774361  1.793244
row_2 -5.033636  9.213437  0.1687509 5.115449 -5.191560

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

      col_1 col_2 col_3 col_4 col_5
row_1  TRUE  TRUE  TRUE  TRUE  TRUE
row_2  TRUE  TRUE  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
WARNING: you appear to work on 32-bit system... large files not supported
WARNING: you appear to work on 32-bit system... large files not supported
unique.names = TRUE
```

```
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames_T
cachesizeMb = 1
number of columns (variables) =  2
number of rows (observations) =  5
usedRowIndex: 1  2  3  4  5
usedColIndex: 1  2
Upper-left 2 columns and  5 rows:
           [,1]       [,2]
[1,] -8.5108500 -5.0336356
[2,] -7.3717060  9.2134371
[3,] -0.5829301  0.1687509
[4,]  6.7743611  5.1154494
[5,]  1.7932440 -5.1915603

> x <- databel_filtered_R("test_matrix_dimnames_T")

WARNING: you appear to work on 32-bit system... large files not supported

> t(testmatr)

           row_1      row_2
col_1 -8.5108500 -5.0336355
col_2 -7.3717062  9.2134366
col_3 -0.5829301  0.1687509
col_4  6.7743612  5.1154496
col_5  1.7932441 -5.1915602

> x

finalizing AbstractMatrix: 0x147f200
finalizing AbstractMatrix: 0x154b000
finalizing AbstractMatrix: 0x14abc00
unique.names = TRUE
unique.rownames = TRUE
unique.colnames = TRUE
backingfilename = test_matrix_dimnames_T
cachesizeMb = 1
number of columns (variables) =  2
number of rows (observations) =  5
usedRowIndex: 1  2  3  4  5
usedColIndex: 1  2
Upper-left 2 columns and  5 rows:
           [,1]       [,2]
[1,] -8.5108500 -5.0336356
[2,] -7.3717060  9.2134371
```

```
[3,] -0.5829301  0.1687509
[4,]  6.7743611  5.1154494
[5,]  1.7932440 -5.1915603
```

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

```
           row_1      row_2
col_1 -8.5108500 -5.0336356
col_2 -7.3717060  9.2134371
col_3 -0.5829301  0.1687509
col_4  6.7743611  5.1154494
col_5  1.7932440 -5.1915603
```

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

```
      row_1 row_2
col_1  TRUE  TRUE
col_2  TRUE  TRUE
col_3  TRUE  TRUE
col_4  TRUE  TRUE
col_5  TRUE  TRUE
```

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