Dates in R

Claudio Sartori

University of Bologna
DISI – Department of Computer Science and Engineering

Dates in R

R doesn't know something is a date or time unless you tell it
(Charlotte Wickham, Oregon State University) #### A quick introduction to *dates* and *times* in R

0.0.1 Date/time classes

Three date/time classes are built-in in R: Date, POSIXct, and POSIXlt.

0.0.2 Date

This is the class to use if you have only dates, but no times, in your data.

[1]:
```
# set the date localisation to english
Sys.setlocale("LC_TIME","C")
```

'C'

[2]:
```
# get current time
Sys.Date()
Sys.time()
```

2021-05-17

```
[1] "2021-05-17 09:35:34 CEST"
```

[3]:
```r
# create a date:
dt1 <- as.Date("2012-07-22")
dt1
```

2012-07-22

[4]:
```r
# next day
dt1 + 1
```

2012-07-23

[5]:
```r
# compare date with string
dt1 == "2012-07-22"
```

TRUE

[6]:
```r
# comparison gives true, but nevertheless
# the string is not a true date
# Executing an expression inside `tryCatch` prevents program stop in case of error
tryCatch({d <- "2012-07-23" + 1} # try the wrong assignment
         , error = function(e) {message("there was an error","\n",e)}
        )
```

```
there was an error
Error in "2012-07-23" + 1: non-numeric argument to binary operator
```

[7]:
```r
# non-standard formats must be specified:
dt2 <- as.Date("04/20/2011", format = "%m/%d/%Y")
dt2 # internal representation and default output do not change
```

2011-04-20

[8]:
```r
dt3 <- as.Date("October 6, 2010", format = "%b %d, %Y")
dt3
```

2010-10-06

Calculations with dates

[9]:
```
# find the difference between dates:
dt1 - dt2
```

```
Time difference of 459 days
```

[10]:
```
difftime(dt1, dt2, units = "weeks")
```

```
Time difference of 65.57143 weeks
```

[11]:
```
# Add or subtract days:
dt2 + 10
```

2011-04-30

[12]:
```
dt2 - 10
```

2011-04-10

[13]:
```
# create a vector of dates and find the intervals between them:
three.dates <- as.Date(c("2010-07-22", "2011-04-20", "2012-10-06"))
three.dates
```

1. 2010-07-22 2. 2011-04-20 3. 2012-10-06

0.0.3 The `diff` function

Given a vector `v` with `n` elements, it computes the differences of the consecutive elements $\{(v_{i+1} - v_i) \forall i = 1, \ldots, n-1\}$

0.0.4 Arguments

- `x` - a numeric vector or matrix containing the values to be differenced.
- `lag` - an integer indicating which lag to use
  - the output is `d[i] = x[i+1]-x[i]`;

- `differences` - an integer indicating the order of the difference
  - 1 for the differences, 2 for the differences of the differences, 3 ...

0.0.5 Details

`diff` is a generic function with a default method and ones for classes "ts", "POSIXt" and "Date"

0.0.6 Value

If `x` is a vector of length `n` and `differences = 1`, then the computed result is equal to the successive differences `x[(1+lag):n] - x[1:(n-lag)]` If difference is larger than one this algorithm is applied recursively to `x`. Note that the returned value is a vector which is shorter than `x`. If `x` is a matrix then the difference operations are carried out on each column separately.

[14]:
```
v <- c(10,20,40)
myDiff <- function(v){
    n <- length(v)
    result <- vector(mode = "double")
    for (i in 1:(n-1)){
        result[i] <- v[i+1] - v[i]
    }
    return(result)
}
```

[15]:
```
diff(v) == myDiff(v)
```

1. TRUE 2. TRUE

[16]:
```
n <- length(v)
diff(v) == v[2:n]-v[1:(n-1)]
```

1. TRUE 2. TRUE

[17]:
```
# Generate a number of random dates and sort them
#    in ascending order (from the earliest to the latest)
set.seed(42)
nDates = 6
```

```
startDate = "2019-01-01"
someDates <- sort(as.Date(startDate) + sample(1000, nDates))
someDates
```

1. 2019-03-16 2. 2019-06-03 3. 2019-08-17 4. 2019-11-18 5. 2020-07-15 6. 2021-09-24

[18]:
```
diff(someDates)
## Time differences in days
```

```
Time differences in days
[1]  79  75  93 240 436
```

[19]:
```
diff(someDates, lag=2)
```

```
Time differences in days
[1] 154 168 333 676
```

[20]:
```
diff(someDates, differences=3)==diff(diff(diff(someDates)))
```

1. TRUE 2. TRUE 3. TRUE

[21]:
```
diff(someDates, differences=2)
```

```
Time differences in days
[1]  -4  18 147 196
```

[22]:
```
d <- 2
x <- someDates
repeat {
    y <- x[-1]
    x <- y - x[-length(x)]
    if (d>1){
        d <- d-1
    } else {
        break
```

```
    }
}
x
```

```
Time differences in days
[1]  -4  18 147 196
```

[23]:
```
someDates[-length(someDates)]
```

1. 2019-03-16 2. 2019-06-03 3. 2019-08-17 4. 2019-11-18 5. 2020-07-15

[34]:
```
# create a sequence of dates:
six.weeks <- seq(dt1, length = 6, by = "week")
six.weeks
```

1. 2012-07-22 2. 2012-07-29 3. 2012-08-05 4. 2012-08-12 5. 2012-08-19 6. 2012-08-26

[35]:
```
six.bi_weeks <- seq(dt1, length = 6, by = 14)
six.bi_weeks
```

1. 2012-07-22 2. 2012-08-05 3. 2012-08-19 4. 2012-09-02 5. 2012-09-16 6. 2012-09-30

[36]:
```
six.bi_weeks <- seq(dt1, length = 6, by = "2 weeks")
six.bi_weeks
```

1. 2012-07-22 2. 2012-08-05 3. 2012-08-19 4. 2012-09-02 5. 2012-09-16 6. 2012-09-30

[39]:
```
# see the internal integer representation
unclass(dt1)
```

15543

[41]:
```
unclass(as.Date("1970-01-01"))
```

0

[43]:
```
dt1 - as.Date("1970-01-01")
```

```
Time difference of 15543 days
```