Nested apply as an alternative to double for loops

MA Alicja Fraś

Poznan University of Economics and Business

16th May 2018

Need to go through columns and apply some specific function on each of them

Feature 1	Feature 2	A	B	C	…
X	1				
Y	2				
Z	2				
⋮	⋮				

1

Mean rate of return within each week of the year for each fund

Date	Week n0	Fund A	Fund B	Fund C	…
2015-01-02	1	0.035	0.034	0.024	
2015-01-05	2	0.013	0.034	-0.012	
2015-01-07	2	-0.012	0.036	0.020	
⋮					

The first solution that came to my mind was…

```
r <- matrix(NA, 105, 1)
for (i in 2:(ncol(data)-1)) {
    r <- cbind(r, as.matrix(aggregate(data[, i], by =
list(data$week.no), FUN = mean)[,2]))
}
```

…but the smarter way would be:

```
apply(data[, 3:ncol(data)], 2, function(x) tapply(x,
data$week.no, FUN = mean))
```

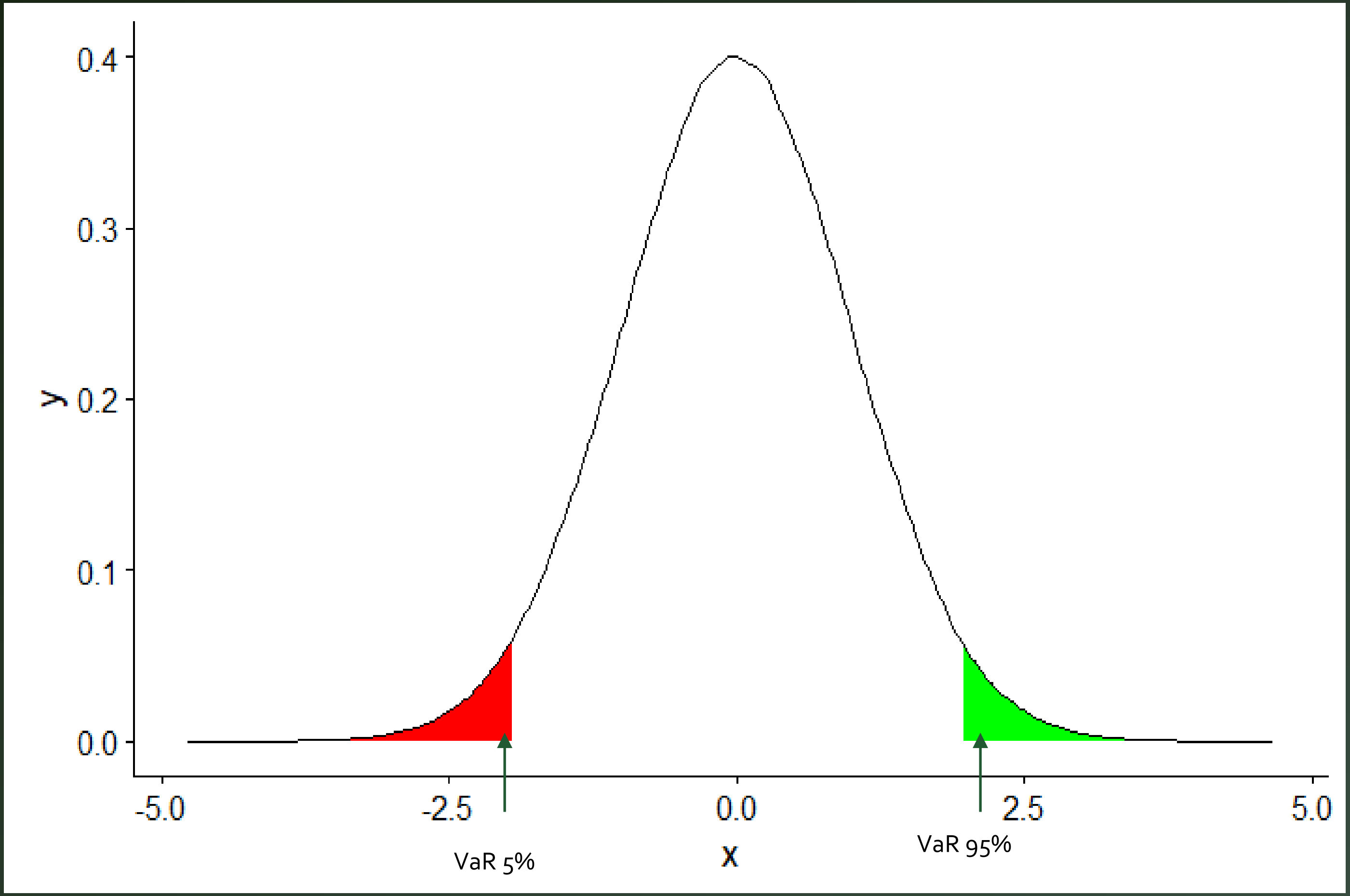
Result

Week no.	Fund A	Fund B	Fund C	…
1	0.019	0.0340	0.024	
2	0.018	0.0338	0.021	
3	0.012	0.0367	0.022	
…				

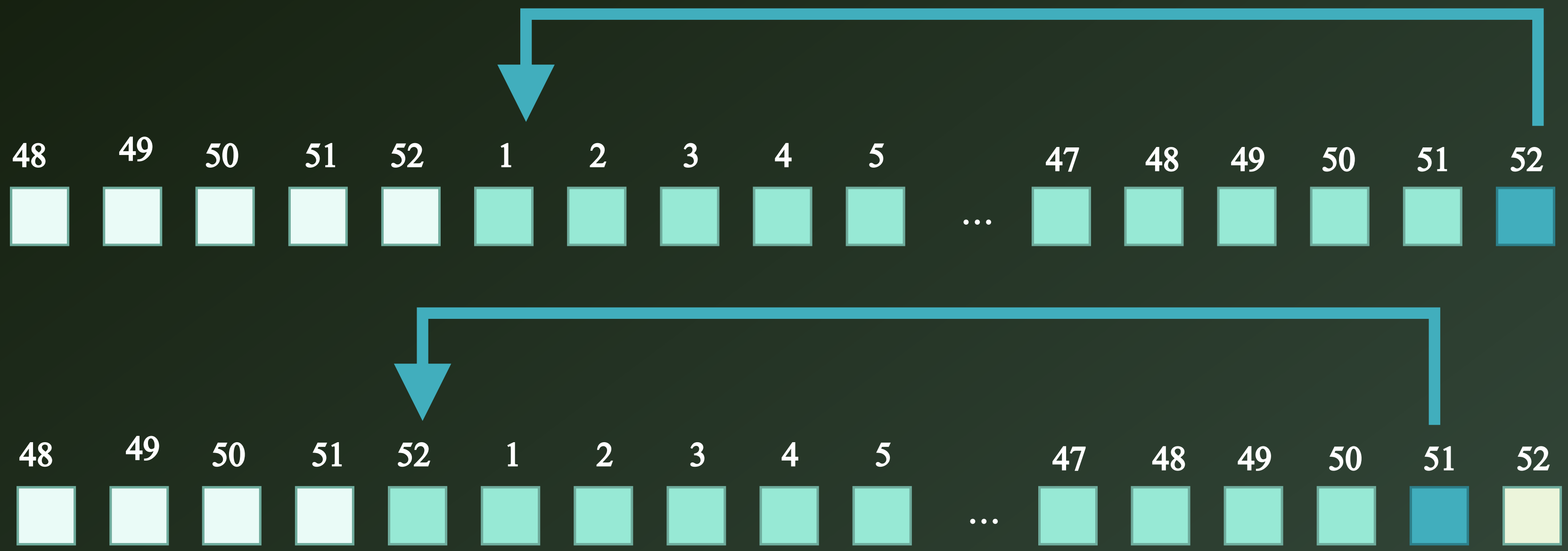
Execution time in seconds (*microbenchmark)*

	Minimum	Mean	Maximum	# of simulations
Double for loop	4.8	5.2	5.7	10
Nested apply	1.8	2.0	2.3	10
Ratio	2.7	2.6	2.5	-

Value at Risk at q% level – boarderline level of loss for given probability (q).



Moving window



Value at Risk with moving time-window: for loop solution…

```
p <- data.frame()
for (i in 1:ncol(data_w)) {
  for (j in 1:52) {
    p[j, i] <- sd(data_w[j:(j+51), i], na.rm = T) *
qnorm(0.05)
  }
}
```

…and nested apply:

```
apply(data_w, 2, function(x)
  sapply(1:52, function(y)
    sd(x[y:(y+51)], na.rm = T) * qnorm(0.05))
```

Execution time measured with *microbenchmark* function [in seconds]

	Minimum	Mean	Maximum	# of simulations
Double for loop	10.4	11.3	12.6	10
Nested apply	2.6	3.4	3.9	10
Ratio	4.0	3.3	3.2	-

Is the fund evaluated daily?

Does it have more than one non-zero rate of return in a week?

Double for loop solution

```
s <- data.frame(rep(NA, 105))
for (i in 2:(ncol(v)-1)) {
  s[,i-1] <- (aggregate(v[,i], by = list(v$week.no.year),
FUN = function(x) sum(!(x == 0))== 1))[,2]
}
```

3

Double apply solution

```
apply(v[,2:(ncol(v)-1)], 2, function(x)
  tapply(x, v$week.no.year, function(y) sum(!(y == 0)) == 1))
```

Execution time measured with *microbenchmark* function [in seconds]

	Minimum	Mean	Maximum	# of simulations
Double for loop	4.0	5.2	8.4	10
Nested apply	1.1	1.6	2.6	10
Ratio	3.6	3.2	3.2	-