

Theory VI: vectorized computation in R

- vectorized computation
- special functions `sweep`, `apply`, `lapply`/`sapply`
- interface to `c`

1

Vectorized computation

(the way not to) Compute inner product of two vectors:

```
> a=rnorm(10000)
> b=rnorm(10000)
> d=rep(0,10000)
> for (i in 1:10000)
+   d[i]=a[i]*b[i]
> s=0
> for (i in 1:10000)
+   s=s+d[i]
> s
[1] -28.97014
```

Vectorized version:

```
> sum(a*b)
[1] -28.97014      Simpler and faster code !
```

2

Another example: simulate unbalanced one way analysis of variance.

```
> nogroups=5
> groupmeans=c(3,4,2,4,5)
> groupsizes=c(10,11,10,9,8)
```

Using for loops:

```
> y=c()
> for (i in 1:nogroups)
+   for (j in 1:groupsizes[i])
+     y=c(y,groupmeans[i]+rnorm(1,sd=0.1))
```

Vectorized version:

```
> y=rep(groupmeans,groupsizes)+rnorm(sum(groupsizes),sd=0.1)
```

3

Computation on matrices and data frames: apply and sweep

Usage: `apply(X,MARGIN,FUN)` `sweep(X,MARGIN,STATS,FUN)`

Compute means of variables NS, ES and area in tortoise

```
> means=apply(tortoise[,c("NS","ES","area")],2,mean)
> means
      NS      ES      area
93.44444 28.81481 290.78148
```

Create data frame with means subtracted:

```
> centered=sweep(tortoise[,c("NS","ES","area")],2,means,"-")
> names(centered)
[1] "NS" "ES" "area"
> mean(centered$NS)
[1] -7.368591e-15 #=0
```

4

Computation on lists and vectors: `lapply` and `sapply`.

Similar but `sapply` seeks to return vector or matrix rather than list:

```
> tortoise.subset=list(tortoise$NS[1:10],tortoise$area[5:21])
> lapply(tortoise.subset,quantile,probs=c(0.025,0.975))
[[1]]
 2.5% 97.5%
2.225 96.100
[[2]]
 2.5% 97.5%
0.078 3055.388
> sapply(tortoise.subset,quantile,probs=c(0.025,0.975))
      [,1]      [,2]
2.5%  2.225  0.078
97.5% 96.100 3055.388
```

5

Calling c-programs from R: if you really have problems with computing time !

- use R CMD SHLIB command for compiling c code containing procedure `compute`.
- use `dyn.load()` to load compiled code.
- interface to c procedure `compute`: `.C("compute",x=...)`.

6