Assignment 9

SiruiZHAO

2023-03-14

Load Library

```
library(doParallel)
library(parallel)
library(foreach)
```

Cores Number

```
#check number of cores
detectCores()
```

```
## [1] 8
```

So, the number of cores in my system is 8.

For Loop

```
#set seed
set.seed(100000)

#create empty vector
means <- numeric(4000)

#start time
start_1 <- Sys.time()

#serial loop
for (i in 1:4000) {
  means[i] <- mean(rnorm(100000, mean = 10, sd = 3))
}

#end time
end_1 <- Sys.time()
```

Time Used

```
RunTime <- end_1 - start_1
RunTime
```

```
## Time difference of 54.22681 secs
```

Multi-threaded program & Serial Program

```
#multi-threading runtime
MTRunTime <- RunTime/8
#how much faster
faster <- RunTime - MTRunTime
faster
```

```
## Time difference of 47.44846 secs
```

As shown above, the run time needed for the serial loop is around 55 seconds for 4000 times. In the serial loop, the program will execute one loop at a time. With parallel programming, we break up repetitive steps and run them at the same time. So, we can use the run time of a serial loop divided by the number of cores (8) to get the theoretical run time for parallel program. Then, we can subtract two runtimes to get the time difference between two programs and this is how much faster a multi-threaded program will run compared to a serial program.

Loop Modify

```
#find core number
Cores <- parallel::makeCluster(detectCores())

#activate Multi-threading
doParallel::registerDoParallel(Cores)

#start time
start_2 <- Sys.time()

#rewrite for loop
parallel <- foreach(i = 1:4000, .combine = 'rbind') %dopar% {
  means[i] <- mean(rnorm(100000, mean = 10, sd = 3))
  }

# end time
end_2 <- Sys.time()

#de-activate multi-threading
parallel::stopCluster(Cores)
```

Run Time

```
#calculate the run time for parallel loop
RunTime_para <- end_2 - start_2
RunTime_para
```

```
## Time difference of 11.09671 secs
```

```
#compare with serial loop and theoretical run time
cat("serial loop run time: ", RunTime, "\n")
```

```
## serial loop run time:  54.22681
```

```
cat("theoretical parallel run time: ", MTRunTime, "\n")
```

```
## theoretical parallel run time:  6.778351
```

```
cat("parallel loop run time: ", RunTime_para)
```

```
## parallel loop run time:  11.09671
```

Results

Compared to the run time of the serial loop, the parallel loop is much more effective. And our actual run time of the parallel loop is close to the theoretical parallel run time, which is around 5~10 seconds. The theoretical parallel run time is calculated from the run time of serial time divided by the core numbers of my computer (8). These close results show that parallelization is effective on this computer.

Thanks!