blblm-vignette

```
library(blblm)
```

#Introduction

“blblm” is bag of little bootstraps for linear regression model. It is a procedure which incorporates features of both the bootstrap and subsampling to yield a robust, computationally efficient means of assessing the quality of estimators (based on linear regression model). For each bootstrap, we sample n from b with replacement instead of sample b from b as in oridinary bootstrap.

##- sample without replacement the sample s times into sizes of b - for each subsample - resample each until sample size is n, r times - compute the bootstrap statistic (e,g., the mean of a variable, or cor between two variables) for each bootstrap sample - compute the statistic (e.g., confidence interval) from the bootstrap statistics - take the average of the statistics

#How to Implement

#First, we need to call blblm function to implement bag of little bootstraps. For example, we want to sample 3 from 100 with replacement.We need to input formula, dataset, m and B. If you want to process with more than one core to get a faster speed, you also need to specify the number of cores you want to use. (nthreads = …)

```
fit <- blblm(mpg ~ wt * hp, data = mtcars, m = 3, B = 100, nthreads = 4)
```

#Print formula

#In case that you want to see the blblm model you are using, you could print is with print.blblm()

```
print.blblm(fit)
```

#Estimation of standard deviation

Using sigma.blb(), you can get the estimated standard deviation. And if you also want to know the 95% confidence interval of the standard deviation, you need to add “confidence = TRUE” in sigma.blb(). Same as above, if you want to have shorter processig time, you need to sepcify the number of cores you want to use with “nthreads = …”. Such as:

```
sigma.blblm(fit, nthreads = 4)
sigma.blblm(fit, confidence = TRUE, nthreads = 4)
```

#Estimation of coefficients

In order to get the coefficients of the model, you need to use coef.blblm().And if you want to get the confidence interval of them, you could use confint.blblm(), to speed up, you can specify the number of cores you want to use. And the default confidence level is 95%, you can change it with “level = …”

```
coef.blblm(fit)
confint.blblm(fit, level = 0.99, nthreads = 4)
```

#Do some Prediction

After building and knowing the blb model, now we can use it to do some prediction. Use predict.blbm(object, new_data, confidence = FALSE, level = 0.95, nthreads = 1) to do prediction with new data. If you also want to get the confidence interval, change “confidence = TRUE”.

```
predict.blblm(fit, data.frame(wt = c(2.5, 3), hp = c(150, 170)), confidence = TRUE)
```