

Estudo de Caso 03: Comparação de desempenho de duas configurações de um algoritmo de otimização

Augusto (Checker), Mateus Pongelupe(Coordinator), Samuel Leite(Recorder)

30 de Outubro de 2018

Resumo

Este relatório é o terceiro dos Estudos de Casos na disciplina de Planejamento e Análise de Experimentos. O problema inicial consiste em duas variações do mesmo algoritmo avaliando a mesma população. Deseja-se analisar o efeito dessa variação afim de verificar o desempenho das alterações no algoritmo.

O método da Evolução Diferencial (DE) é uma técnica que procura solucionar problemas de otimização contínua. É um método simples, porém poderoso, que avalia soluções de forma iterativa, buscando otimização global. Ele é dotado de 3 passos, sendo eles: mutação, cruzamento e seleção. No primeiro, é feita a soma de um indivíduo base com a diferença de dois outros indivíduos aleatórios multiplicados por uma ponderação F . O vetor mutado é chamado de V . O cruzamento é uma intercalagem entre o vetor original e o vetor mutado, de forma que aumenta-se a variabilidade da sua população. O vetor intercalado é chamado de U . A seleção avalia qual o melhor indivíduo a ser escolhido, seja ele da população original ou da população cruzada.

Planejamento do Experimento

Será realizado um experimento, nos quais parâmetros diferentes são submetidos ao realizar o método DE. Para o teste de hipóteses, avaliada a diferença entre as médias, denominada pelo parâmetro μ_d .

$$\begin{cases} H_0 : \mu_D = 0 \\ H_1 : \mu_D \neq 0 \end{cases}$$

Para esse teste, definiu-se um nível de significância de $\alpha = 0.05$, uma mínima diferença de importância prática $d^* = \delta^*/\sigma = 0.5$ e uma potência desejada de $\pi = 1 - \beta = 0.8$.

O tamanho da amostra é calculado utilizando-se dos parâmetros de teste requeridos e a função *power.t.test*.

Coleta de Dados

Coleta de Dados

Para geração dos dados foi criada a rotina **generateAllSamples** que recebe o número de execuções de cada algoritmo por amostra e o número de amostras N calculado para o problema. Nessa função, são selecionadas aleatoriamente as dimensões das funções de Rosenbrock a serem avaliadas e os resultados são escritos em arquivos no formato *.csv*. Dessa forma, essa função é executada apenas uma vez para um dado tamanho amostral. O valor para o número de avaliações por software será setado no valor padrão de 30.

```
library("smoof")

suppressPackageStartupMessages(library(smoof))

getSamples <- function(n,dim) {
  # This assingment makes fn a global variable
  # This function is created dinamically for each dimension
```

```

fn <- function(X){
  if(!is.matrix(X)) X <- matrix(X, nrow = 1) # <- if a single vector is passed as X
  Y <- apply(X, MARGIN = 1,
            FUN = smoof::makeRosenbrockFunction(dimensions = dim))
  return(Y)
}

selpars <- list(name = "selection_standard")
stopcrit <- list(names = "stop_maxeval", maxevals = 5000 * dim, maxiter = 100 * dim)
probpars <- list(name = "fn", xmin = rep(-5, dim), xmax = rep(10, dim))
popsize = 5 * dim

## Config 1
recpars1 <- list(name = "recombination_lbga")
mutpars1 <- list(name = "mutation_rand", f = 4.5)

## Config 2
recpars2 <- list(name = "recombination_blxAlphaBeta", alpha = 0.1, beta = 0.4)
mutpars2 <- list(name = "mutation_rand", f = 3)

library("ExpDE")

suppressPackageStartupMessages(library(ExpDE))

# Extracted one execution of the DE algorithm
generate_sample <- function(mutpars, recpars, popsize, selpars, stopcrit, probpars) {
  return(ExpDE(mutpars = mutpars,
               recpars = recpars,
               popsize = popsize,
               selpars = selpars,
               stopcrit = stopcrit,
               probpars = probpars,
               showpars = list(show.iters = "dots", showevery = 20))$Fbest);
}

# Executes n times a given configuration of the DE algorithm
generate_n_samples <- function(n, mutpars, recpars) {
  return(replicate(n, generate_sample(mutpars, recpars, popsize, selpars, stopcrit, probpars)))
}

# Runs n times both DE algorithms
return(data.frame("A1" = generate_n_samples(n, mutpars1, recpars1), "A2" = generate_n_samples(n, mutp
}

generateAllSamples <- function(n_samples, N) {
  rosenbrok_dim_interval <- 2:150
  dimensions <- sort(sample(rosenbrok_dim_interval, N))

  result <- lapply(dimensions, FUN = function(x) {
    write.csv(getSamples(n_samples, x), paste('samples-dim-', x, '.csv', sep = ''), row.names=FALSE)
  })
  write.csv(data.frame("dim" = dimensions), 'dimensions.csv', row.names=FALSE)
}

```

```

}

dimensionsFile <- 'dimensions.csv'
nTestsPerSample <- 30

# if dimension files does not exist, gets all samples
if(!file.exists(dimensionsFile)) {
  generateAllSamples(nTestsPerSample, N)
}

```

Como os dados foram gerados apenas uma vez, devido ao tempo requerido para tal, o processamento dos dados foi feito a partir da leitura dos arquivos *.csv* produzidos. Assim, todos os dados salvos para experimento, estão presentes na variável *dataFrame*.

```

dimensions <- read.csv(file = dimensionsFile, header = T)
dataFrame <- data.frame('A1' = character(), 'A2' = character(), 'dim' = character(), stringsAsFactors =
lapply(dimensions$dim, FUN = function(x) {
  fileName <- paste('samples-dim-', x, '.csv', sep = '')
  data <- read.csv(file = fileName, header = T)
  dataFrame <-- rbind(dataFrame, data)
}))

```

Análise Estatística

Teste da Média

Dados os parâmetros definidos na seção *Planejamento do Experimento* para o teste, foram analisadas $N = 34$ amostras e o teste foi executado nas linhas abaixo. Foram utilizados valores diferentes para a análise DE, de acordo com o fornecido para o Grupo E.

Avaliando suposições do modelo

Conclusões e Recomendações

O teste realizado neste relatório serviu para avaliar o desempenho de dois algoritmos difentes avaliando uma mesma população. O método da Evolução Diferencial é um método de otimização, mas como ele apresenta parâmetros numéricos de entrada, deve também ter esses otimizados. Este relatório teve como objetivo analisar dois conjuntos parâmetros para realizar o método de convergencia DE para um mesmo conjunto de dados gerados e observar qual obteve uma maior otimização.

Referências

- R Man Pages - asbio package - <https://rdrr.io/cran/asbio/man/power.z.test.html>
- R Man Pages - car package - <https://rdrr.io/cran/car/man/qcPlot.html>
- Statistics R Tutorial - <https://www.cyclismo.org/tutorial/R/confidence.html>
- Montgomery, Douglas C. - Applied statistics and probabiliy for engineers (3? Edi??o) - Cap?tulos 8,9
- Notas de Aula - <https://github.com/fcampelo/Design-and-Analysis-of-Experiments>
- Notas - https://edisciplinas.usp.br/pluginfile.php/2063723/mod_resource/content/0/Aula11-2016.pdf