

Trabalho Final - Avaliação de Algoritmos em um Problema de Otimização de Roteiros de Viagem

Equipe 04

03 de Julho de 2017

Coordenador: Gustavo Vieira
Relator: Bernardo Marques
Verificador: Danny Tonidandel
Monitor: Alessandro Cardoso

Sumário

...

1. Descrição do Problema

As ferramentas atuais para planejamento de viagens permitem que usuários busquem passagens e hotéis e construam seu próprio itinerário, mas não oferecem suporte para viagens com flexibilidade nas datas ou na ordenação de destinos. Para solucionar esse problema, foi proposto um novo sistema que busca as melhores combinações de passagens e hotéis para uma viagem, minimizando o custo e tempo de transporte. O sistema busca otimizar roteiros com datas flexíveis e não requer uma ordem fixa dos destinos, de forma que considere todas as possíveis configurações que atendem as especificações do usuário. Dessa forma, oferece boas opções de itinerários para viagens com esforço do usuário muito menor, representando uma melhoria significativa na experiência de planejamento de viagens.

O conjunto de restrições do problema considera datas inicial e final, datas específicas para cada destino, ordenação dos destinos e características das acomodações e transportes. Pode ser considerado uma aplicação multiobjetivo do problema do caixeiro viajante com janelas de tempo e dependência temporal.

Para realizar a otimização de roteiros, foi utilizado um algoritmo de colônia de formigas [?]. Uma melhoria para o algoritmo utilizando busca local foi proposta. O objetivo do trabalho é comparar os algoritmos e determinar se a melhoria proposta é melhor para solucionar o problema, avaliando tanto o tempo gasto na otimização e a qualidade do conjunto de soluções encontrado. Assim, as perguntas a serem respondidas são:

- TODO
- TODO

No entanto, o tamanho do roteiro a ser avaliado também influencia esses resultados. Assim, é necessário considerar o tamanho da instância como uma covariável do problema. Para lidar com isso, serão utilizadas técnicas de blocagem, e também exploradas técnicas estatísticas de análise de covariância.

Os parâmetros experimentais desejados são:

- Nível de significância: $\alpha = 0.05$;
- Tamanho de efeito de interesse prático: $d_{time}^* = d_{quality}^* = 1$;
- Potência desejada: $(1 - \beta) = \pi \geq 0.8$.

TODO: explicar melhor o problema

2. Planejamento Experimental

Os dados experimentais do problema foram obtidos através de simulações. Foram geradas instâncias do problema de diferentes tamanhos. Para cada uma delas, cada algoritmo foi aplicado 8 vezes, e os resultados registrados. A ordem das instâncias geradas e da aplicação dos algoritmos foi aleatorizada de maneira a bloquear possíveis efeitos de ordem.

Nos experimentos, foi medido o tempo gasto na execução de cada algoritmo. Para minimizar a influência de outros processos do sistema, a execução foi limitada a um único núcleo do processador, com prioridade de execução alta.

Cada algoritmo de otimização produz como resultado um conjunto de soluções. A qualidade desse conjunto é medida através do Hipervolume [?] do conjunto na fronteira de Pareto, considerando como referência o ponto da solução Nadir na fronteira de Pareto (isto é, a pior solução teórica). Para cada instância avaliada, um novo ponto de referência é calculado através de uma heurística.

O método de Hipervolume é bastante sensível ao ponto de referência definido [?]. Sendo assim, a qualidade medida entre diferentes instâncias não tem significado quando comparadas entre si. No entanto, a métrica é eficiente para comparar a qualidade de diferentes conjuntos de solução encontrados no mesmo ponto. Sendo assim, é necessário avaliar a qualidade relativa das soluções em cada instância, e não o seu valor absoluto.

2.1 Definição do Tamanho Amostral

O tamanho amostral do problema foi calculado a partir dos requisitos experimentais estabelecidos. Uma vez que o tamanho de efeito de interesse prático foi definido em relação a variância, não era necessário conhecer a variância dos dados previamente. Assume-se que a variância entre os diferentes algoritmos é igual.

Dessa forma, o tamanho amostral do experimento foi calculado através da equação:

$$n = 2 \left(\frac{1}{d_t^*} \right)^2 (t_{\alpha/2} + t_\beta)^2,$$

, solucionada iterativamente.

2.2 Análise de Variância com Blocagem

Para a avaliar existência de diferenças significativas entre os dois parâmetros, será utilizado o teste estatístico ANOVA.

Essa técnica analisa médias e variâncias de observações de diferentes grupos para verificar se existe diferença estatística significativa entre as médias desses grupos. Assim, generaliza o teste t para mais de dois grupos, permitindo que sejam comparados simultaneamente [Campelo, 2015], [Montgomery and Runger, 2011]. Em outras palavras, a análise de variância é utilizada quando se quer decidir se os níveis apresentam médias diferentes em relação a uma média global μ . As diferenças entre as médias dos níveis e a média global é $\tau_i \forall i$.

Vale ressaltar que a técnica é restrita apenas à indicação de existência ou não de diferenças entre os níveis avaliados, sem indicar quais níveis seriam diferentes. Além disso, quando a análise de variância tem como resultado um indicativo de refutação da hipótese nula (1) é que podem ser evidenciados os indícios de diferenças entre os níveis.

$$\begin{cases} H_0 : \tau_i = 0, & \forall i \\ H_1 : \exists \tau_i \neq 0 \end{cases} \quad (1)$$

Esta etapa foi considerada como uma primeira verificação ao questionamento deste estudo por apresentar uma abordagem que pode minimizar o custo de coleta de amostras menor do que o teste de comparações

múltiplas discutido na Seção 2.2. Portanto, caso não seja detectada alguma diferenças entre os níveis, as comparações múltiplas são dispensáveis, o que reduz o custo do experimento.

2.3 Comparações Múltiplas

Caso a análise ANOVA identifique a existência de diferenças entre os parâmetros analisados nos algoritmos, deve-se proceder com testes de comparação múltiplas. As comparações serão realizadas em relação ao algoritmo *ACO*, de maneira a verificar se a outra alternativa proposta é melhor. Desta forma, caso a análise de variância indique a existência de diferenças entre os níveis, será aplicado o teste de comparações múltiplas um-contra-todos (one-vs-all) de *Dunnnett*, onde o *ACO com busca local* será confrontado com o algoritmo *ACO*.

Para realizar o teste de comparações múltiplas, deve-se manter o controle sobre os erros do tipo-*I* em cada comparação de maneira com que ele não se acumule a cada teste sucessivo. Assim, os valores de α são corrigidos para cada teste através do método de correção de Bonferroni (2):

$$\alpha_{adj} = \frac{\alpha_{família}}{K}, \quad (2)$$

no qual $K = 1$ consiste no número de comparações a serem feitas que no caso do teste um-contra-todos de *Dunnnett*. O número de comparações nesse caso é dado pela Equação (3):

$$K = a - 1; \quad (3)$$

2.4 Análise de Covariância (ANCOVA)

A análise de covariância é uma técnica – assim como a blocagem ou pareamento (em testes-t) – bastante útil para a melhora da precisão de um experimento [Montgomery, 1984]. Em diversos aspectos é similar à análise de variância (ANOVA), porém permite ter controle sobre a influência do covariante nas variáveis dependentes.

A covariável complementa o controle local e, obviamente, necessita estar correlacionada com a variável de resposta para que se possa fazer uso de tal análise. E quando a análise de variância é realizada com uma ou mais covariáveis, é usual chamar a análise de ANCOVA.

Na ANCOVA a variável dependente é contínua (e.g. tempo, velocidade, etc.), enquanto que a variável independente é normalmente categórica (e.g. “masculino/feminino”, “fumante/não-fumante” etc.). A análise de variância, por sua vez, poderá converter-se na ANCOVA quando for adicionado um covariante, que consiste em outra variável, que pode ser tanto categórica quanto contínua.

Existem aliás, duas razões principais para sair-se da ANOVA e passar a ser considerada a análise de covariância, i.e., razões técnicas para adicionar um covariante, quais sejam:

- Reduzir fatores de variabilidade inter-grupos, ou alcançar um nível maior de entendimento a partir da variância desconhecida;
- Isolar o efeito que ocorre quando o controle experimental não permite que o experimentador elimine, de maneira razoável, explicações alternativas para uma relação observada entre variáveis independentes e dependentes, o que é chamado de “confusão” *confounding*. Em outras palavras, esta variável é algo que pode estar influenciando o experimento mas que não está, a princípio, no modelo original, e é uma potencial fonte de vieses no experimento.

A ANCOVA, permite, portanto, um controle do erro experimental, aumentando sua precisão. Vale ressaltar que a técnica não é restrita apenas à indicação de existência ou não de diferenças entre os níveis avaliados.

Table 1: ANCOVA como um “modelo ajustado” da ANOVA.

Fontes de variação	\sum de quadrados	Graus de Lib	média quadrática	F_0
regressão	(S_{xy}/S_{xx})	1		
Tratamentos	$\frac{SS'_E - SS_E}{(S_{xy})^2/S_{xx} - [E_{yy} - (E_{xy})^2/E_{xx}]}$	$a - 1$	$\frac{SS'_E - SS_E}{a-1}$	$\frac{1}{MS_E} \frac{SS'_E - SS_E}{a-1}$
Erro	$SS_E = E_{yy} - (E_{xy})^2/E_{xx}$	$a(n-1) - 1$	$\frac{SS_E}{a(n-1)-1}$	
Total	S_{yy}	$an - 1$		

Considerando um experimento com um fator e uma covariável, o modelo estatístico da análise de covariância pode ser estabelecido como [Montgomery and Runger, 2011]:

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \epsilon_{ij}, \quad (4)$$

em que: μ é uma constante, τ_i é o efeito do i -ésimo “tratamento”, X_{ij} é o valor observado da covariável e $\bar{x}_{..}$ a média dos valores x_{ij} (da covariável). Além disso, β consiste no coeficiente de regressão linear entre a covariável X e a variável de resposta Y , com $\beta \neq 0$ e ϵ_{ij} é uma componente aleatória de erro, com $i = 1, \dots, a$ e $j = 1, \dots, n$. Neste caso a relação deve ser linear. Vale ressaltar que este modelo pressupõe que a variável de resposta e a covariável estão relacionadas linearmente. Para efeito de simplificação, contudo, a análise de covariância pode ser vista como uma análise “ajustada” de variância, e pode ser sumarizada na tabela (??):

em que:

$$S_{yy} = \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{y_{..}^2}{an}, \quad (5)$$

$$S_{xx} = \sum_{i=1}^a \sum_{j=1}^n x_{ij}^2 - \frac{x_{..}^2}{an}, \quad (6)$$

$$S_{xy} = \sum_{i=1}^a \sum_{j=1}^n x_{ij}y_{ij} - \frac{(x_{..})(y_{..})}{an}, \quad (7)$$

$$T_{yy} = \frac{1}{n} \sum_{i=1}^a y_{i.}^2 - \frac{y_{..}^2}{an}, \quad (8)$$

$$T_{xx} = \frac{1}{n} \sum_{i=1}^a x_{i.}^2 - \frac{x_{..}^2}{an}, \quad (9)$$

$$T_{xy} = \frac{1}{n} \sum_{i=1}^a x_{i.}y_{i.} - \frac{(x_{..})(y_{..})}{an}, \quad (10)$$

$$E_{yy} = S_{yy} - T_{yy}, \quad (11)$$

$$E_{xx} = S_{xx} - T_{xx}, \quad (12)$$

$$E_{xy} = S_{xy} - T_{xy}. \quad (13)$$

Assim, rejeita-se a hipótese nula $H_0 : \beta = 0$ se $F_0 > F_{\alpha, 1, a(n-1)-1}$.

3. Resultados Análise Estatística

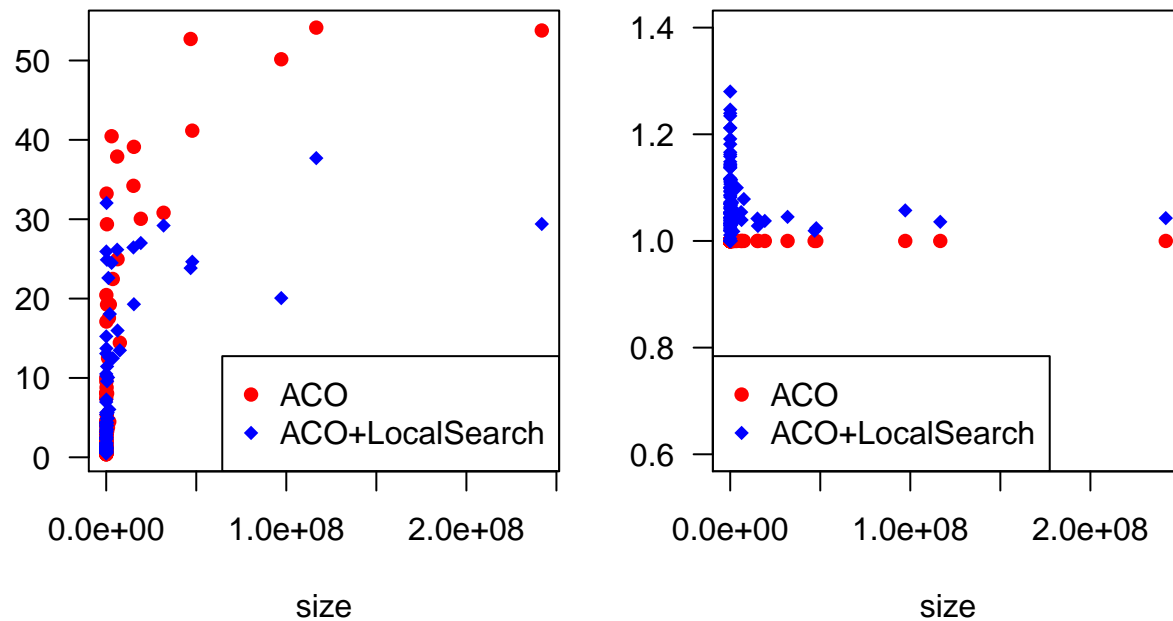
3.1 Tratamento e Validação dos Dados

Considerando o experimento realizado, foi criada uma rotina para validação dos dados obtidos e identificação de erros, onde o tempo de execução e valor de qualidade das respostas dos algoritmos na escala (((XXXXXlogaritma))) devem ser maior que 0.

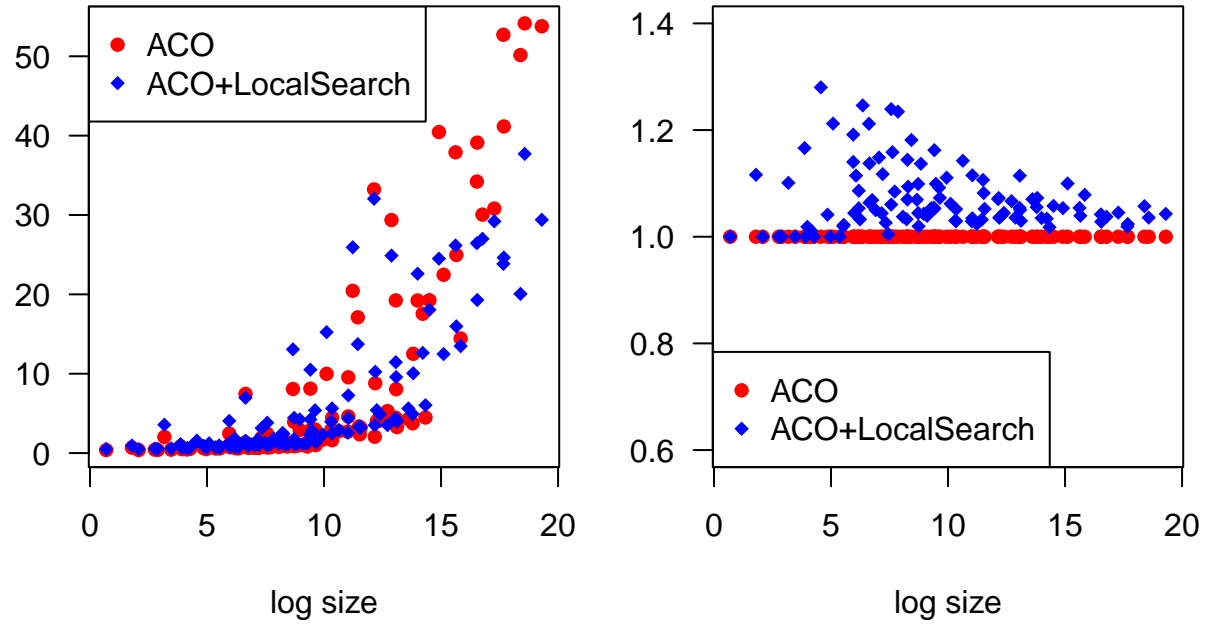
1. Tempo > 0
2. 0 > Qualidade > 1

Caso os valores de uma execução não atendam essas condições, ela seria descartada. No entanto, nenhuma das amostras apresentou tal problema.

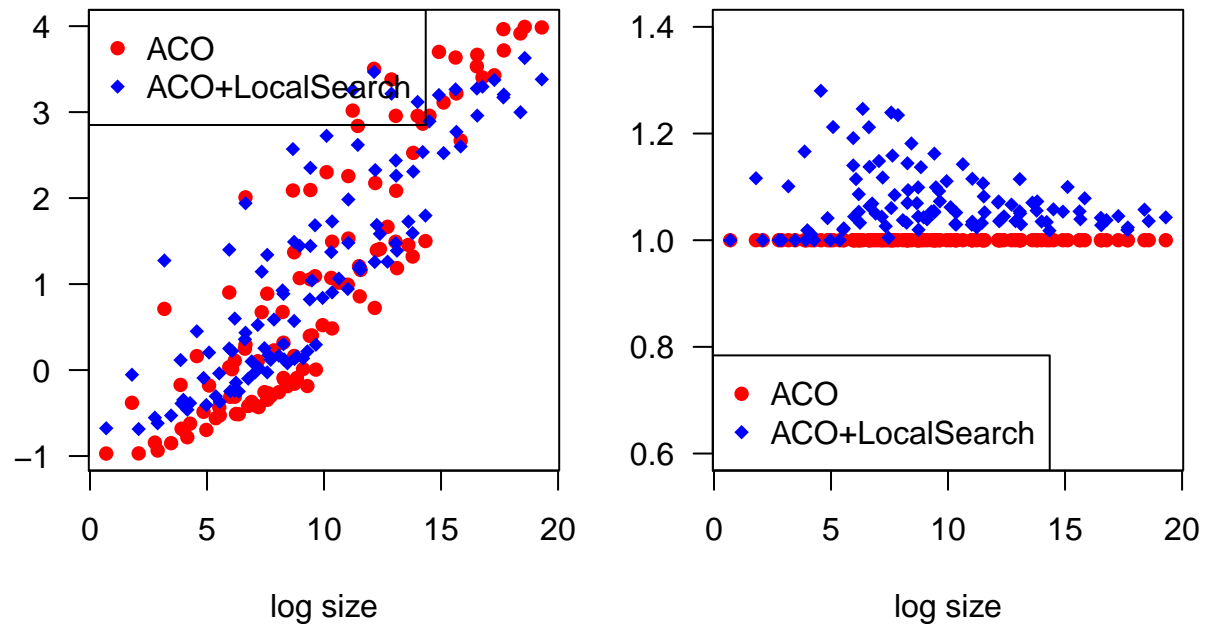
3.2 Análise Exploratória dos Dados



A escala não permite ter uma visualização adequada dos dados. Assim, é aplicada uma transformação logarítmica na variável Size para solucionar esse problema.



Para normalizar a distribuição dos dados e obter uma relação linear entre a covariável e o tempo, foi aplicada uma transformação normal na variável de resposta Time.



Os resultados parecem indicar uma relação linear entre tempo e tamanho de problema (ambos em escala logarítmica). Além disso, a qualidade relativa das soluções parece não depender do tamanho. Finalmente, há indícios de que o algoritmo proposto é melhor que a alternativa tanto em quesitos de tempo quanto qualidade da solução encontrada.

3.2 Análise de Variância com Blocagem

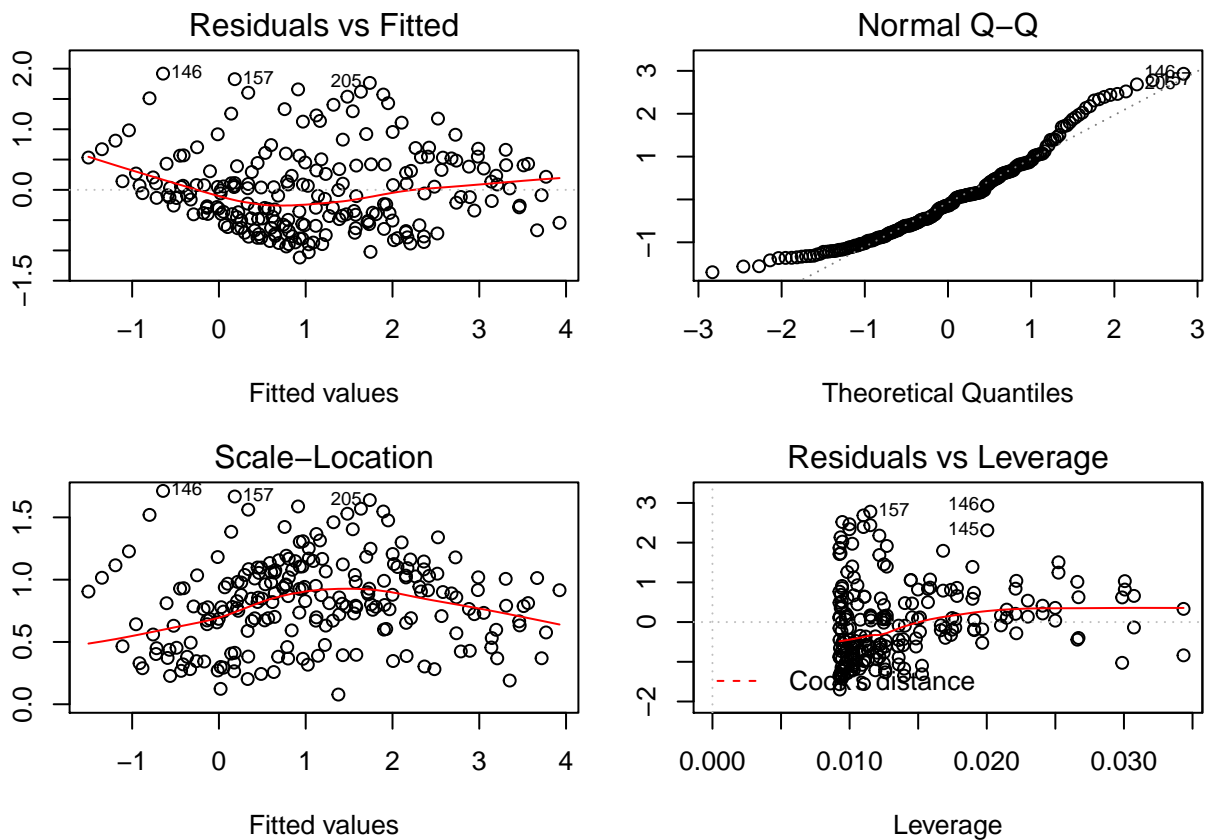
Tempo

Análise inicial sem considerar efeitos de tamanho do problema:

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm    1    1.3    1.308    0.706  0.402
## Residuals  214   396.6    1.853
```

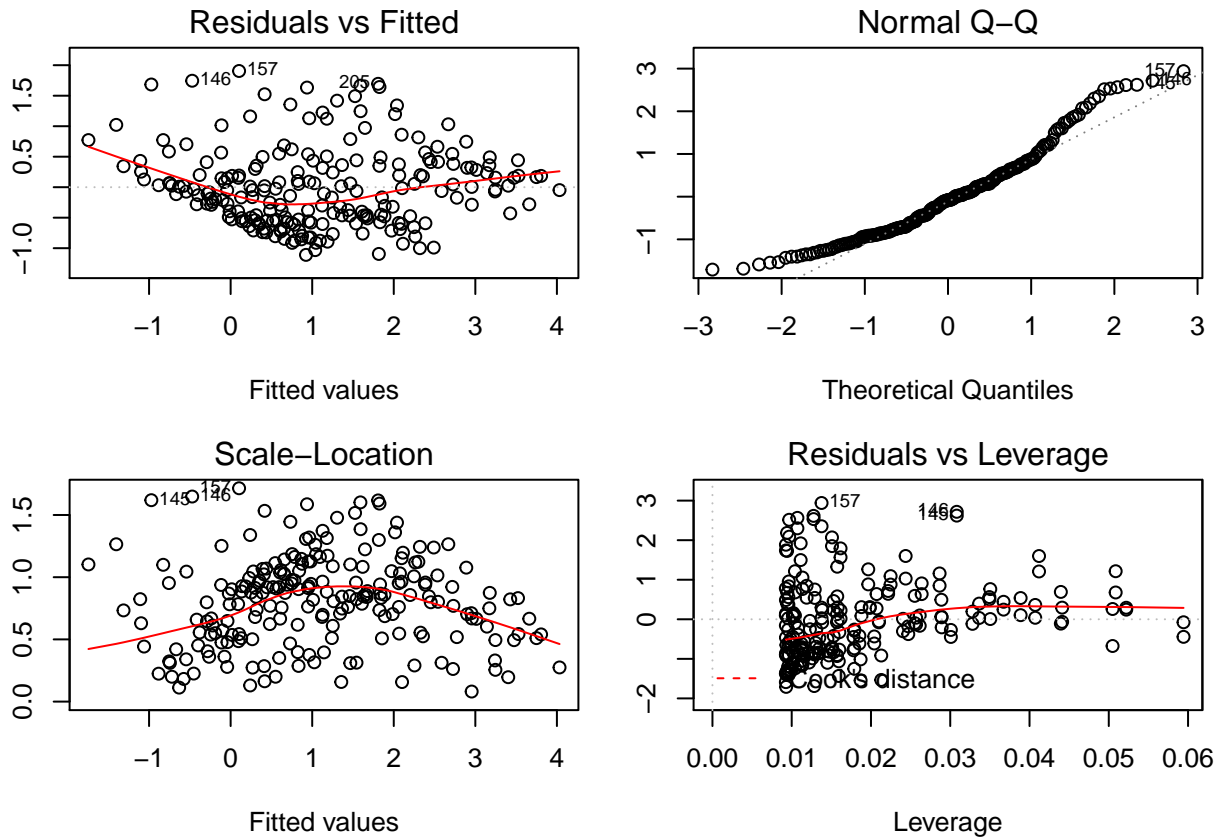
Considerando efeitos:

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm    1    1.31    1.31    2.995  0.085 .
## log(size)    1 303.60 303.60 695.041 <2e-16 ***
## Residuals  213   93.04    0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



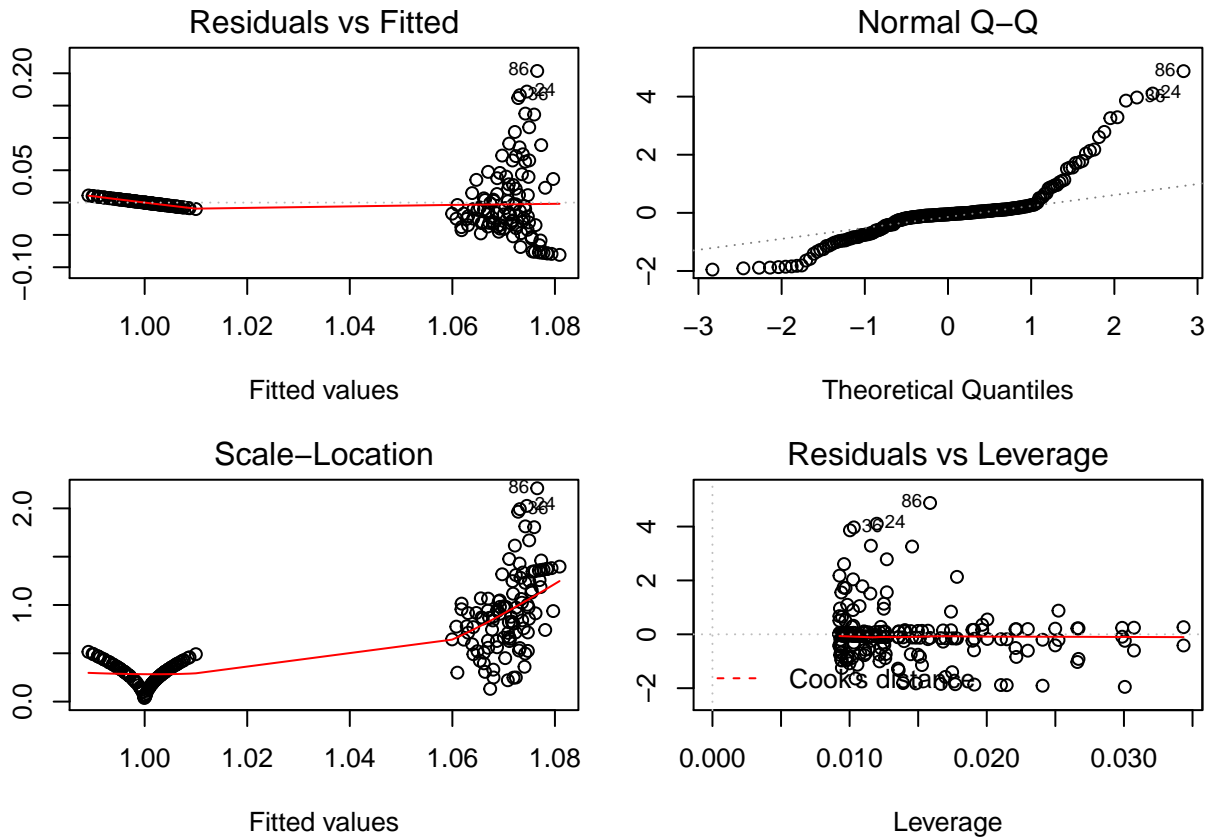
Considerando variável de interação entre tempo e tamanho do problema

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm    1    1.31    1.31    3.072  0.0811 .
## log(size)    1 303.60 303.60 712.948 <2e-16 ***
## algorithm:log(size) 1    2.76    2.76    6.488  0.0116 *
## Residuals  212   90.28    0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Qualidade da Solução

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm    1  0.2716  0.27159 153.372 <2e-16 ***
## log(size)    1  0.0048  0.00479   2.705  0.101
## Residuals  213  0.3772  0.00177
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Paired t-test
##
## data: qualityRelative by algorithm
## t = -12.335, df = 107, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.08231522 -0.05952095
## sample estimates:
## mean of the differences
## -0.07091809
```

3.3 Análise de Covariância

Geração do Conjunto de Dados

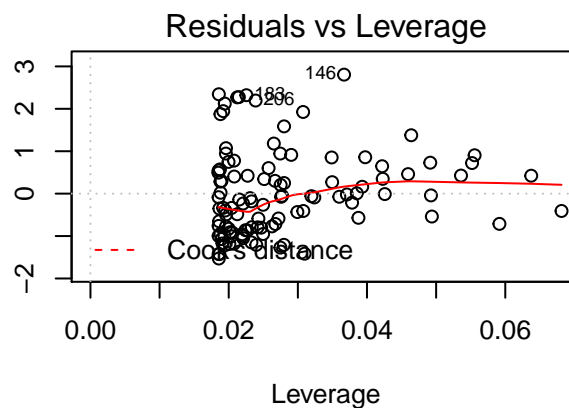
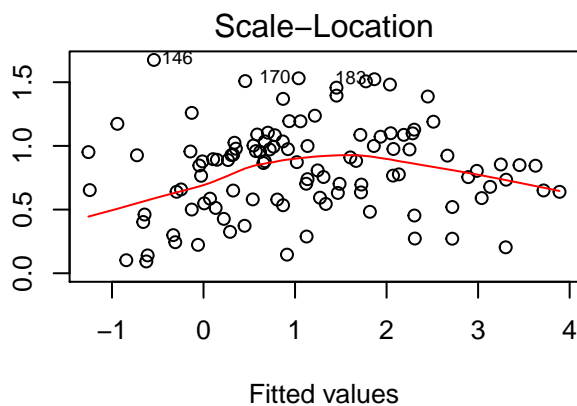
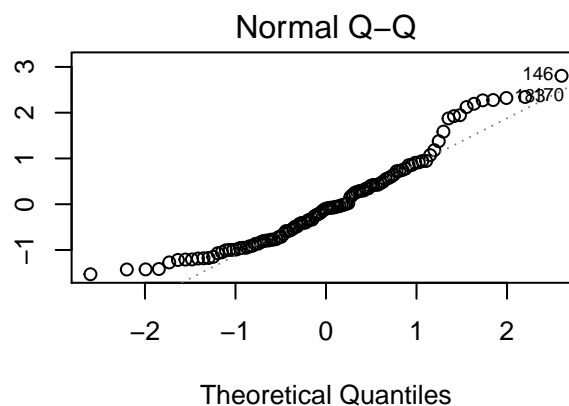
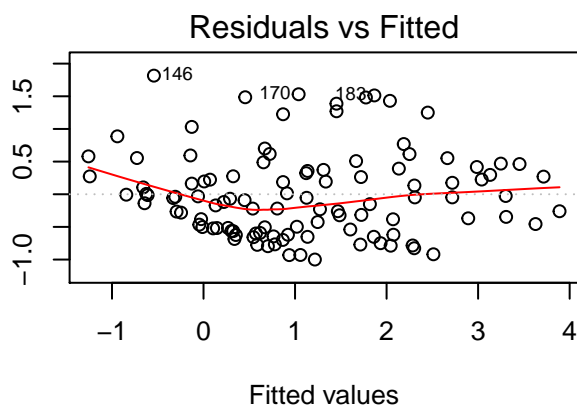
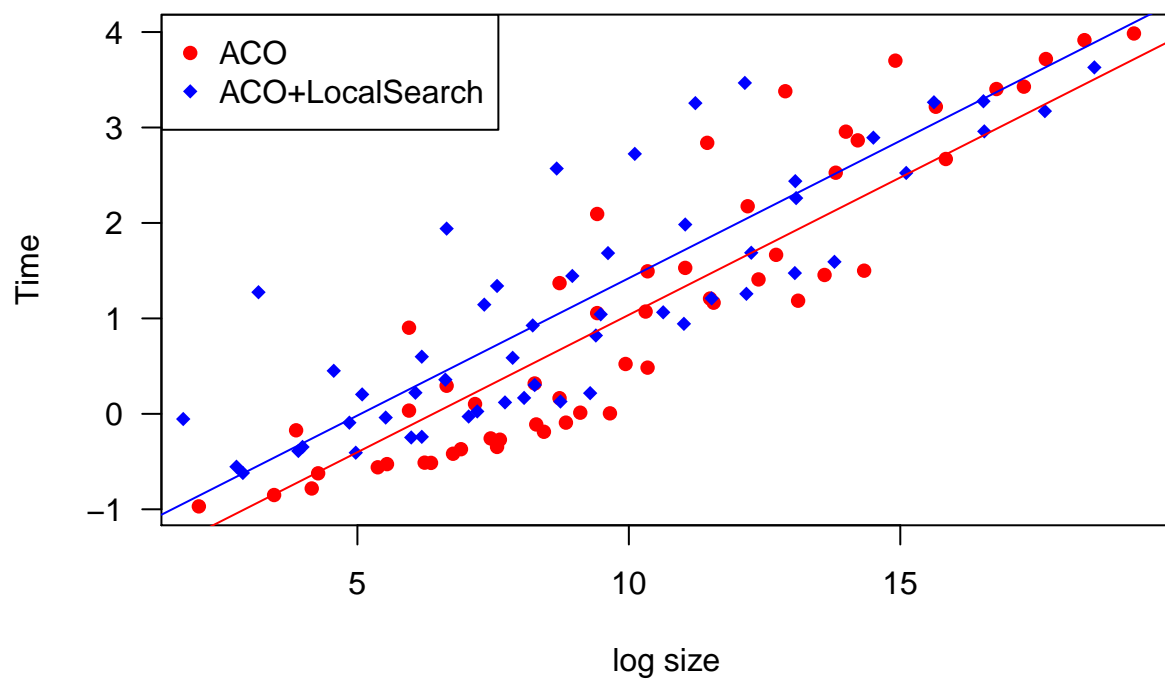
TODO: plots do novo conjunto

Tempo

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm    1   0.07    0.07   0.154  0.695
## log(size)    1 153.64  153.64 353.305 <2e-16 ***
```

```
## Residuals    105  45.66    0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

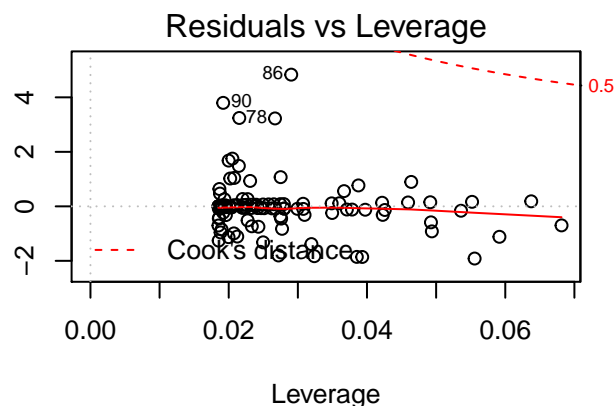
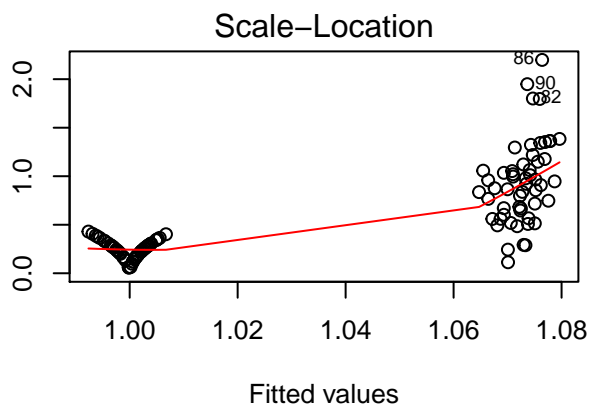
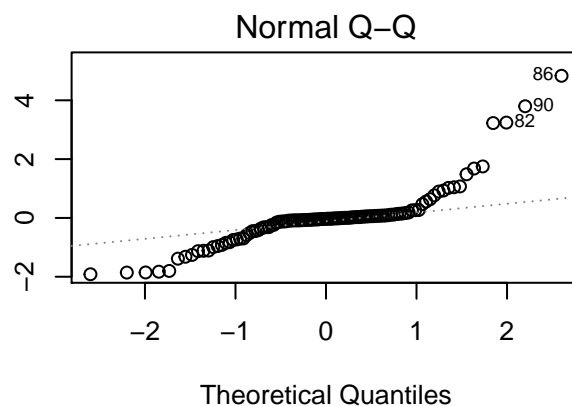
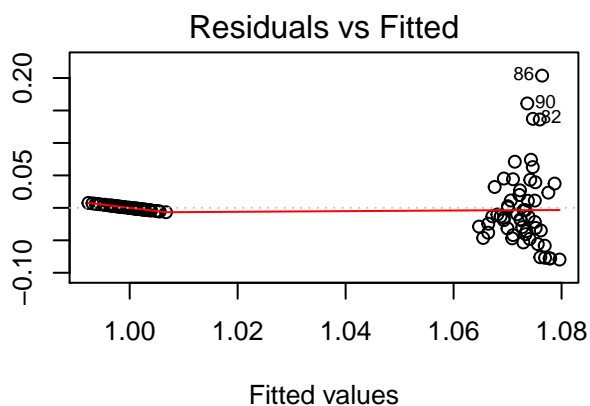
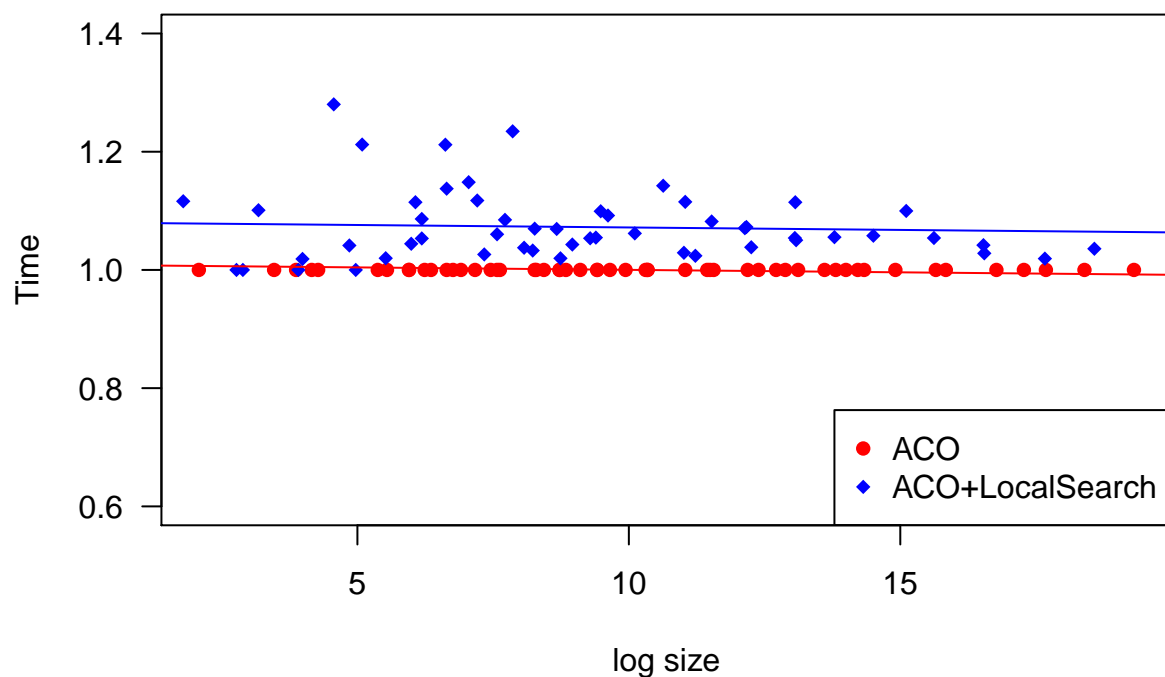
##              Df Sum Sq Mean Sq F value    Pr(>F)
## log(size)      1 149.79   149.79 344.452 < 2e-16 ***
## algorithm      1   3.92    3.92   9.007 0.00336 **
## Residuals     105  45.66    0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Qualidade

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## algorithm      1 0.14276 0.14276   78.049 2.5e-14 ***
## log(size)       1 0.00129 0.00129    0.705  0.403
## Residuals    105 0.19205 0.00183
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##           Df Sum Sq Mean Sq F value Pr(>F)
## log(size)       1 0.00776 0.00776     4.24  0.0419 *
## algorithm      1 0.13629 0.13629   74.51 7.05e-14 ***
## Residuals    105 0.19205 0.00183
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



4. Discussão e Conclusões

Referências

- Felipe Campelo. Lecture notes on design and analysis of experiments. <https://github.com/fcampelo/Design-and-Analysis-of-Experiments>, 2015. Version 2.11, Chapter 7; Creative Commons BY-NC-SA 4.0.
- D. C. Montgomery. *Design and analysis of experiments*, volume 7. Wiley New York, 1984.
- D. C. Montgomery and G. C. Runger. *Applied Statistics and Probability for Engineers*, volume 5. John Wiley and Sons, 2011.