

02

Validação gráfica do modelo

Toda análise realizada para ajustar um modelo linear simples é baseado em várias suposições sobre o modelo, independência, homoscedasticidade e normalidade dos erros.

Uma das maneiras de avaliar o cumprimento das suposições do modelo, é mediante os respectivos gráficos das suposições, lembrando que estas análises gráficas não são confirmatórias, elas são exploratórias para termos uma ideia do que acontece com os resíduos, que são as estimativas dos erros do modelo.

Considere os dados e o modelo das notas de Matemática dos alunos em função das horas de estudo semanais dedicadas à matéria.

```
dados <- read.csv2("~/data.xlsx", sep="")  
modelo = lm(nota ~ hora_estudo, data = dados)
```

Os seguintes comandos servem para obter os gráficos que permitem verificar as suposições do modelo.

Independência dos resíduos:

```
plot(modelo$residuals,  
main = "Resíduos",  
ylab = "")
```

Homoscedasticidade dos resíduos:

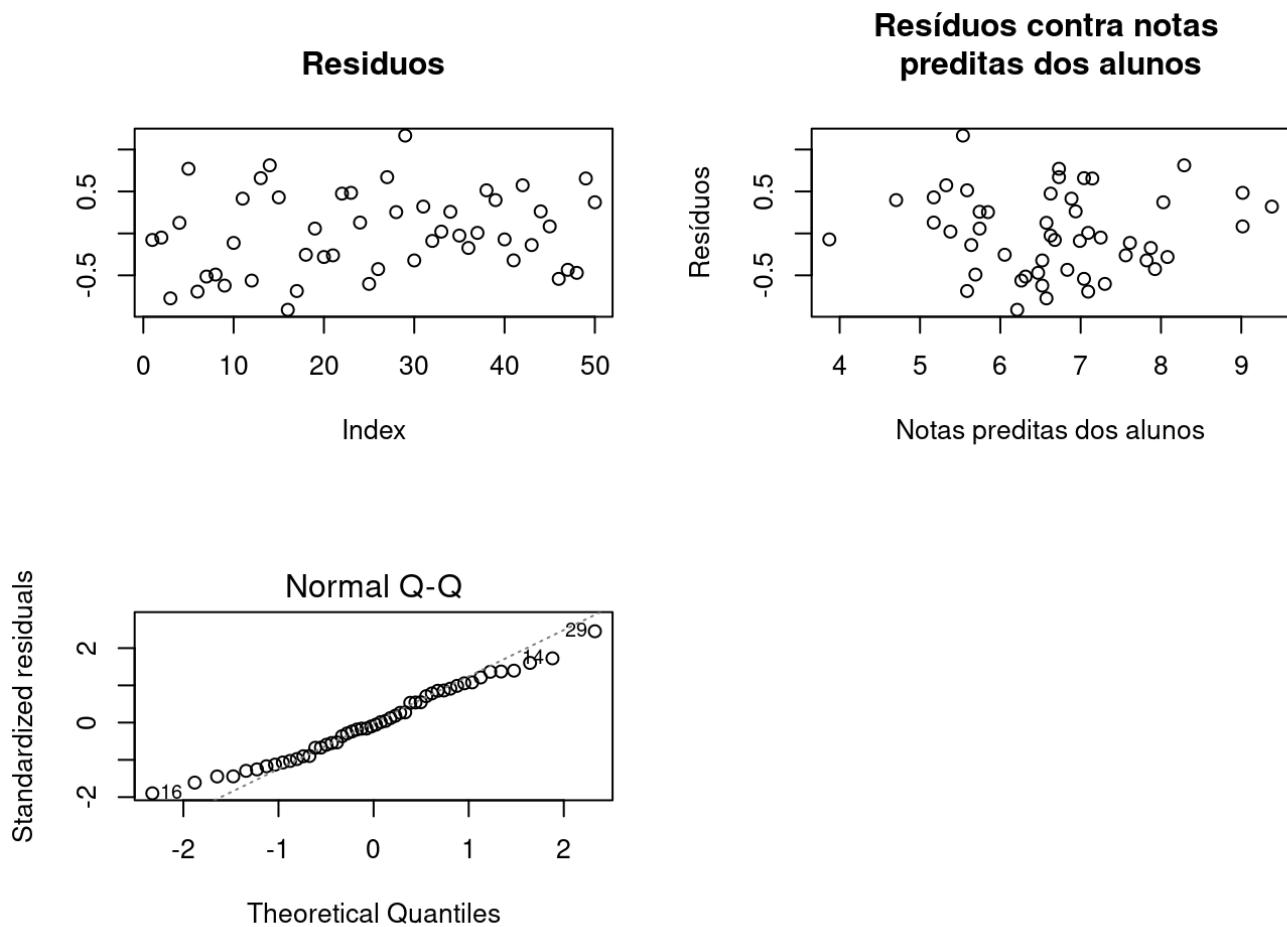
```
plot(modelo$fitted.values, modelo$residuals,  
main = "Resíduos contra notas \n preditas dos alunos \n",  
xlab = "Notas preditas dos alunos",  
ylab = "Resíduos")
```

Normalidade dos resíduos:

```
plot(modelo, 2)
```

Para obter os três gráficos juntos, podemos adicionar antes dos gráficos o comando `par(mfrow = c(2,2))`, que criar uma única figura de tamanho 2x2, com os 3 gráficos:

```
par(mfrow = c(2,2))  
plot(modelo$residuals,  
main = "Resíduos",  
ylab = "")  
plot(modelo$fitted.values, modelo$residuals,  
main = "Resíduos contra notas \n preditas dos alunos \n",  
xlab = "Notas preditas dos alunos",  
ylab = "Resíduos")  
plot(modelo, 2)
```



Lembremos que as análises gráficas não confirmam nem negam o cumprimento das suposições do modelo.

Como vimos em aula, os resíduos não devem mostrar tendência ou padrão algum. Na figura dos Resíduos, podemos verificar que estão bem espalhados, o que indica independência dos resíduos.

O gráfico dos resíduos contra os valores ajustados ou preditos das notas dos alunos também não apresentam relação. Logo, há indícios de homogeneidade dos resíduos.

Finalmente, o gráfico de normalidade Q-Q plot parece seguir a linha de normalidade esperada, sendo que apenas os alunos 16, 14 e 29 se afastam um pouco da normalidade, mas não sabemos se são suficientes para violar esta suposição pois os gráficos são análises exploratórias.

Sempre é recomendável realizar os testes correspondentes para as validações das suposições, ainda mais se temos suspeita da violação destas já que estes gráficos só ajudam para termos uma ideia do que acontece.