asian carp first part

Eddie Wu

2023-06-14

Introduction

This .Rmd file is to show the progress on asian carp temperature and condition analyses. Since sub-sampling from spatial autocorrelation does not give significantly different results from normal analysis, we present the results without sub-sampling here.

For each species, we have four different models:

1. Simple linear model (same slope, same intercept)
2. Linear additive model (same slope, different intercept)
3. Interaction model (different slope, same intercept)
4. Group-specific model (different slope, different intercept)

And we consider two temperature metrics:

1. Annual temperature
2. Winter temperature (temperature from the coldest quarter)

```
library(ggplot2)
library(ggfortify)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
## Import data
asian.carp <- read.csv("asian carp final.csv")
asian.carp$Condition <- as.factor(asian.carp$Condition)

Black <- read.csv("eddie_carp_new.csv")
Black$condition <- as.factor(Black$condition)

## Separate by species
Grass <- asian.carp[asian.carp$Species=="Grass",]
Bighead <- asian.carp[asian.carp$Species=="Bighead",]
Silver <- asian.carp[asian.carp$Species=="Silver",]
Big.sil <- rbind(Bighead, Silver) # combine the two groups
```

Asian carp

Using annual temperature

```
# Clean data
asian.carp.clean <- asian.carp %>%
  filter(Condition %in% c("natural", "artificial"))

# Build the models
asian.simple <- lm(log(AAM)~AnnualTemp, data = asian.carp.clean)
asian.linear <- lm(log(AAM)~AnnualTemp+Condition, data = asian.carp.clean)
asian.int <- lm(log(AAM)~AnnualTemp:Condition, data = asian.carp.clean)
asian.group <- lm(log(AAM)~AnnualTemp*Condition, data = asian.carp.clean)

# Compare the AICs
AIC(asian.simple, asian.linear, asian.int, asian.group)
```

```
##              df      AIC
## asian.simple  3 19.08105
## asian.linear  4 21.05438
## asian.int     4 21.05838
## asian.group   5 23.05379
```

For Asian carp using annual air temperature, the simple linear model works the best.

Using cold temperature

```
# Build the models
asian.simple <- lm(log(AAM)~ColdTemp, data = asian.carp.clean)
asian.linear <- lm(log(AAM)~ColdTemp+Condition, data = asian.carp.clean)
asian.int <- lm(log(AAM)~ColdTemp:Condition, data = asian.carp.clean)
asian.group <- lm(log(AAM)~ColdTemp*Condition, data = asian.carp.clean)

# Compare the AICs
AIC(asian.simple, asian.linear, asian.int, asian.group)
```

```
##              df      AIC
## asian.simple  3 25.05163
## asian.linear  4 26.72525
## asian.int     4 27.03839
## asian.group   5 28.72257
```

Same conclusion when using the cold temperature. But using annual temperature seems to have a better model as explained by lower AIC values.

Grass carp

Using annual temperature

```
# Clean data
grass.clean <- Grass %>%
  filter(Condition %in% c("natural", "artificial"))

# Build the models
grass.simple <- lm(log(AAM)~AnnualTemp, data = grass.clean)
grass.linear <- lm(log(AAM)~AnnualTemp+Condition, data = grass.clean)
```

```r
grass.int <- lm(log(AAM)~AnnualTemp:Condition, data = grass.clean)
grass.group <- lm(log(AAM)~AnnualTemp*Condition, data = grass.clean)

# Compare the AICs
AIC(grass.simple, grass.linear, grass.int, grass.group)
```

```
##              df       AIC
## grass.simple  3  9.848057
## grass.linear  4 10.801064
## grass.int     4 10.759725
## grass.group   5 12.663058
```

For grass carp using annual air temperature, the simple linear model works the best. Although there does not seem to be much difference among the models.

Using cold temperature

```r
# Build the models
grass.simple <- lm(log(AAM)~ColdTemp, data = grass.clean)
grass.linear <- lm(log(AAM)~ColdTemp+Condition, data = grass.clean)
grass.int <- lm(log(AAM)~ColdTemp:Condition, data = grass.clean)
grass.group <- lm(log(AAM)~ColdTemp*Condition, data = grass.clean)

# Compare the AICs
AIC(grass.simple, grass.linear, grass.int, grass.group)
```

```
##              df       AIC
## grass.simple  3  9.573623
## grass.linear  4 11.500450
## grass.int     4 11.407066
## grass.group   5 13.293477
```

Same as using annual temperature.

Bighead and silver carp

Using annual temperature

```r
# Clean data
big.sil.clean <- Big.sil %>%
  filter(Condition %in% c("natural", "artificial"))

# Build the models
big.sil.simple <- lm(log(AAM)~AnnualTemp, data = big.sil.clean)
big.sil.linear <- lm(log(AAM)~AnnualTemp+Condition, data = big.sil.clean)
big.sil.int <- lm(log(AAM)~AnnualTemp:Condition, data = big.sil.clean)
big.sil.group <- lm(log(AAM)~AnnualTemp*Condition, data = big.sil.clean)

# Compare the AICs
AIC(big.sil.simple, big.sil.linear, big.sil.int, big.sil.group)
```

```
##                df      AIC
## big.sil.simple  3 13.82776
## big.sil.linear  4 14.04192
## big.sil.int     4 14.42420
```

```
## big.sil.group   5 16.04189
```

For grass carp using annual air temperature, the simple linear model works the best. Although there does not seem to be much difference among the models.

Using cold temperature

```
# Build the models
big.sil.simple <- lm(log(AAM)~ColdTemp, data = big.sil.clean)
big.sil.linear <- lm(log(AAM)~ColdTemp+Condition, data = big.sil.clean)
big.sil.int <- lm(log(AAM)~ColdTemp:Condition, data = big.sil.clean)
big.sil.group <- lm(log(AAM)~ColdTemp*Condition, data = big.sil.clean)

# Compare the AICs
AIC(big.sil.simple, big.sil.linear, big.sil.int, big.sil.group)
```

```
##                df      AIC
## big.sil.simple  3 18.40785
## big.sil.linear  4 19.10025
## big.sil.int     4 20.39045
## big.sil.group   5 21.09175
```

Same conclusion as annual temperature. Annual temperature fits better models.

Black carp

Using annual temperature

```
# Clean data
black.clean <- Black %>% filter(!row_number() == 15) %>% filter(sex != "male")

# Build the models
black.simple <- lm(log(AAM)~AnnualTemp, data = black.clean)
black.linear <- lm(log(AAM)~AnnualTemp+condition, data = black.clean)
black.int <- lm(log(AAM)~AnnualTemp:condition, data = black.clean)
black.group <- lm(log(AAM)~AnnualTemp*condition, data = black.clean)

# Compare the AICs
AIC(black.simple, black.linear, black.int, black.group)
```

```
##              df         AIC
## black.simple  3  0.03963796
## black.linear  4 -0.93938475
## black.int     4 -1.07498134
## black.group   5  0.76248019
```

The interaction model seems to perfrom slightly better than the others.

Using cold temperature

```
# Build the models
black.simple <- lm(log(AAM)~ColdTemp, data = black.clean)
black.linear <- lm(log(AAM)~ColdTemp+condition, data = black.clean)
black.int <- lm(log(AAM)~ColdTemp:condition, data = black.clean)
black.group <- lm(log(AAM)~ColdTemp*condition, data = black.clean)
```

```r
# Compare the AICs
AIC(black.simple, black.linear, black.int, black.group)
```

```
##              df         AIC
## black.simple  3 -0.53666717
## black.linear  4 -1.73129584
## black.int     4  1.35778275
## black.group   5 -0.07443596
```

The linear additive model seems to be the best one, while the interaction model is the worst fit.

For black carp, if we remove the Southern Ukraine data point:

Using annual temperature

```r
# Remove the Southern Ukraine data point
black.clean.r <- black.clean[black.clean$AAM != 11,]

# Remove the Nucet farm data point
black.clean.r <- black.clean.r[black.clean$AAM != 4,]

# Build the models
black.simple.r <- lm(log(AAM)~AnnualTemp, data = black.clean.r)
black.linear.r <- lm(log(AAM)~AnnualTemp+condition, data = black.clean.r)
black.int.r <- lm(log(AAM)~AnnualTemp:condition, data = black.clean.r)
black.group.r <- lm(log(AAM)~AnnualTemp*condition, data = black.clean.r)

# Compare the AICs
AIC(black.simple.r, black.linear.r, black.int.r, black.group.r)
```

```
##                df       AIC
## black.simple.r  3 -12.94229
## black.linear.r  4 -18.29560
## black.int.r     4 -18.77144
## black.group.r   5 -17.25131
```

Linear addictive and interaction model.

Using annual temperature

```r
# Build the models
black.simple.r <- lm(log(AAM)~ColdTemp, data = black.clean.r)
black.linear.r <- lm(log(AAM)~ColdTemp+condition, data = black.clean.r)
black.int.r <- lm(log(AAM)~ColdTemp:condition, data = black.clean.r)
black.group.r <- lm(log(AAM)~ColdTemp*condition, data = black.clean.r)

# Compare the AICs
AIC(black.simple.r, black.linear.r, black.int.r, black.group.r)
```

```
##                df       AIC
## black.simple.r  3 -14.56695
## black.linear.r  4 -20.05756
## black.int.r     4 -13.06562
## black.group.r   5 -19.46013
```

Linear additive model.

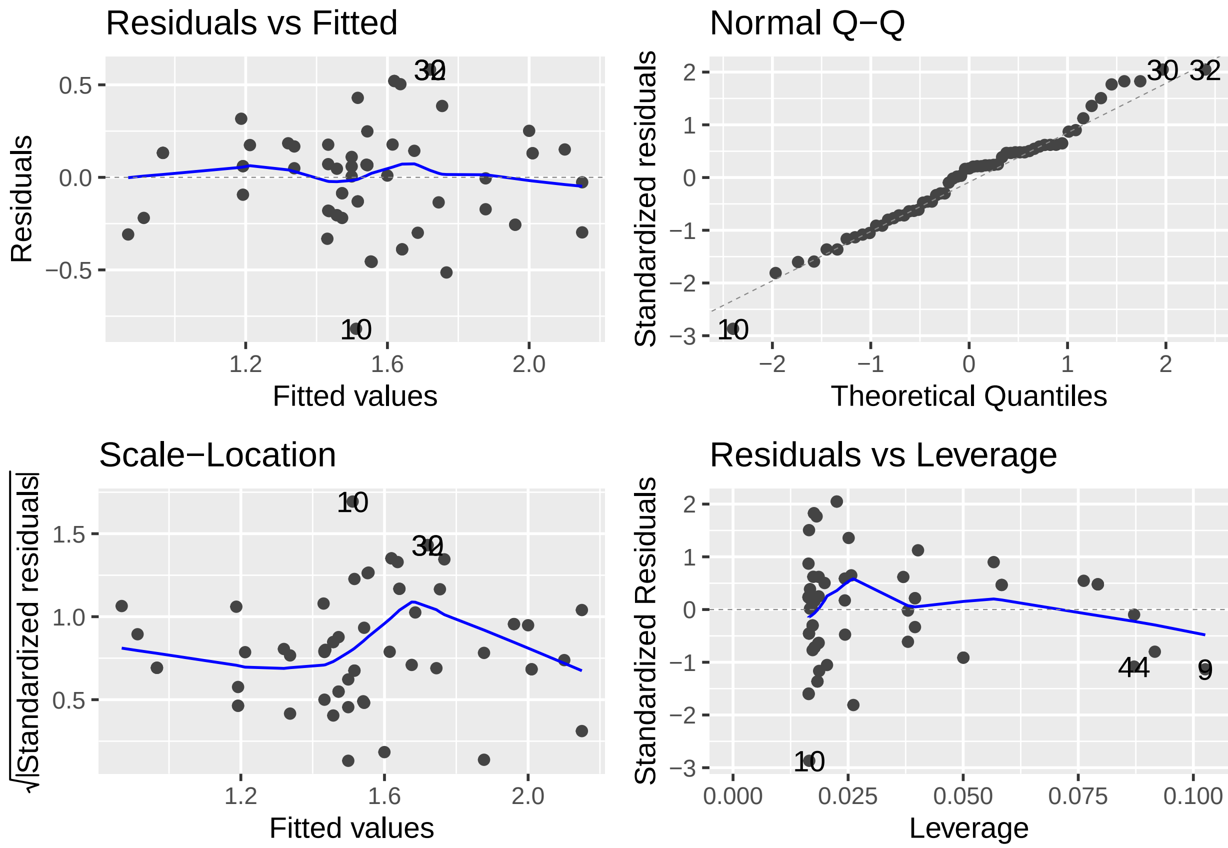
Concluding points

- For all the other asian carp species, the simple linear model seems to be the best fit. Besides, using cold temperature and annual temperature does not produce different results.
- For black carp, the linear additive model (*same slope, different intercept*) seems to be a better fit. Using cold temperature on the original data suggests that the interaction model is the best (probably because the winter temperature at these two locations are very similar.)
- After removing the two data points, we see that the linear additive still seems to be a better fit.

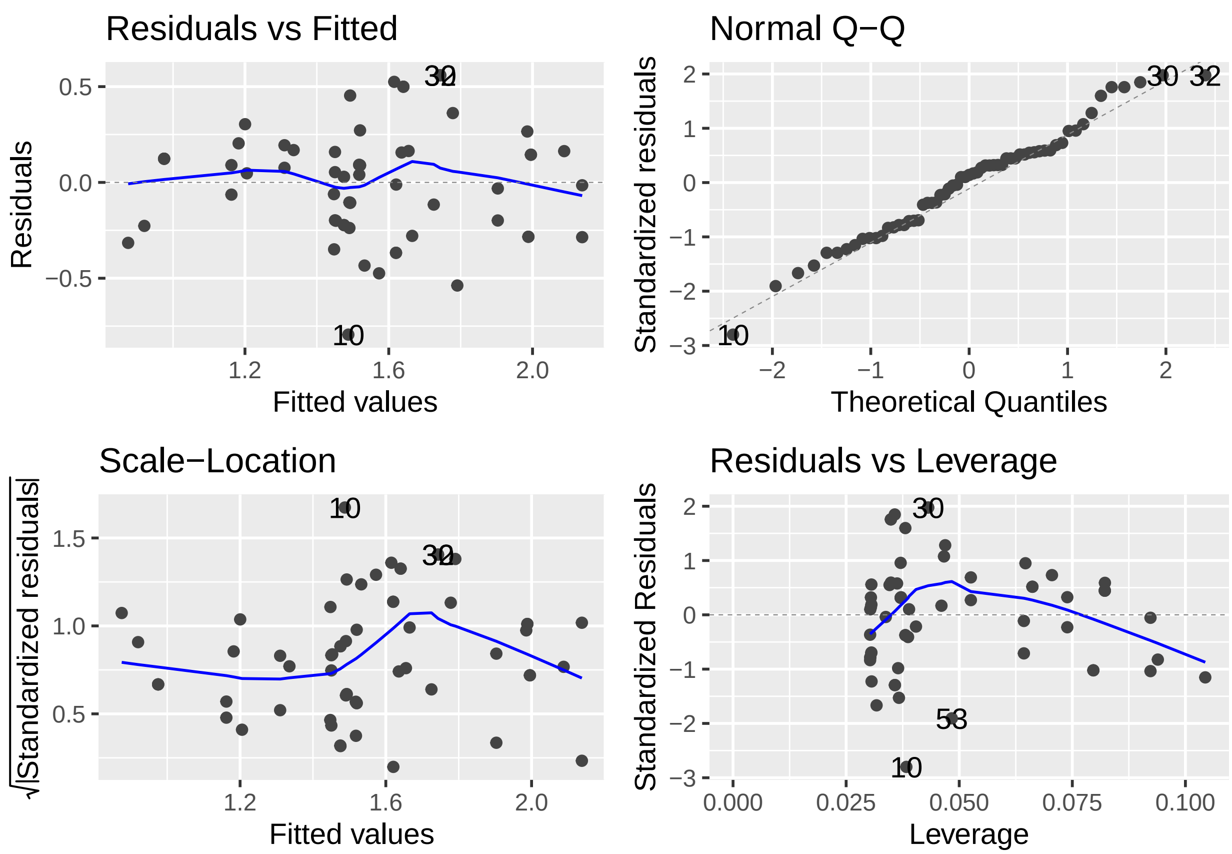
Diagnostic plots

Only the diagnostic plots with cold temperature

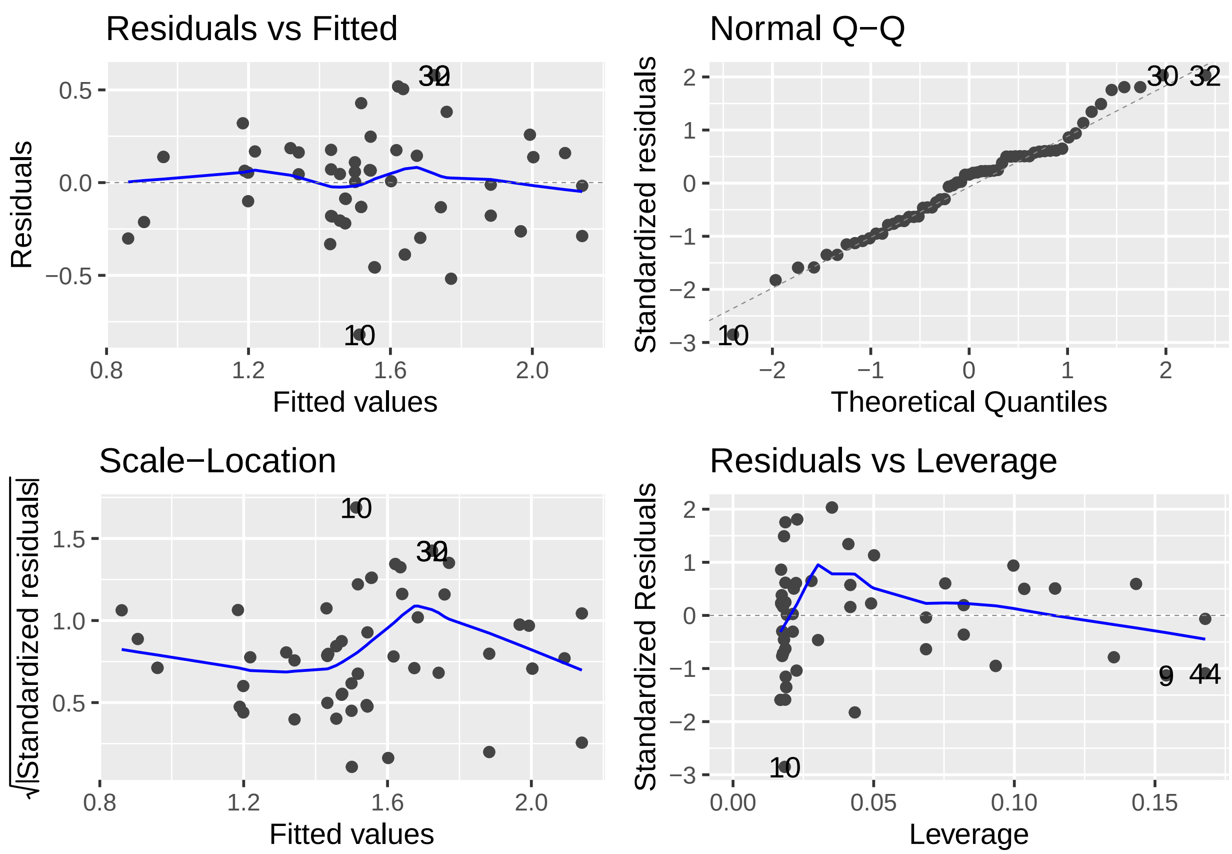
```
# Asian carp

autoplot(asian.simple)
```



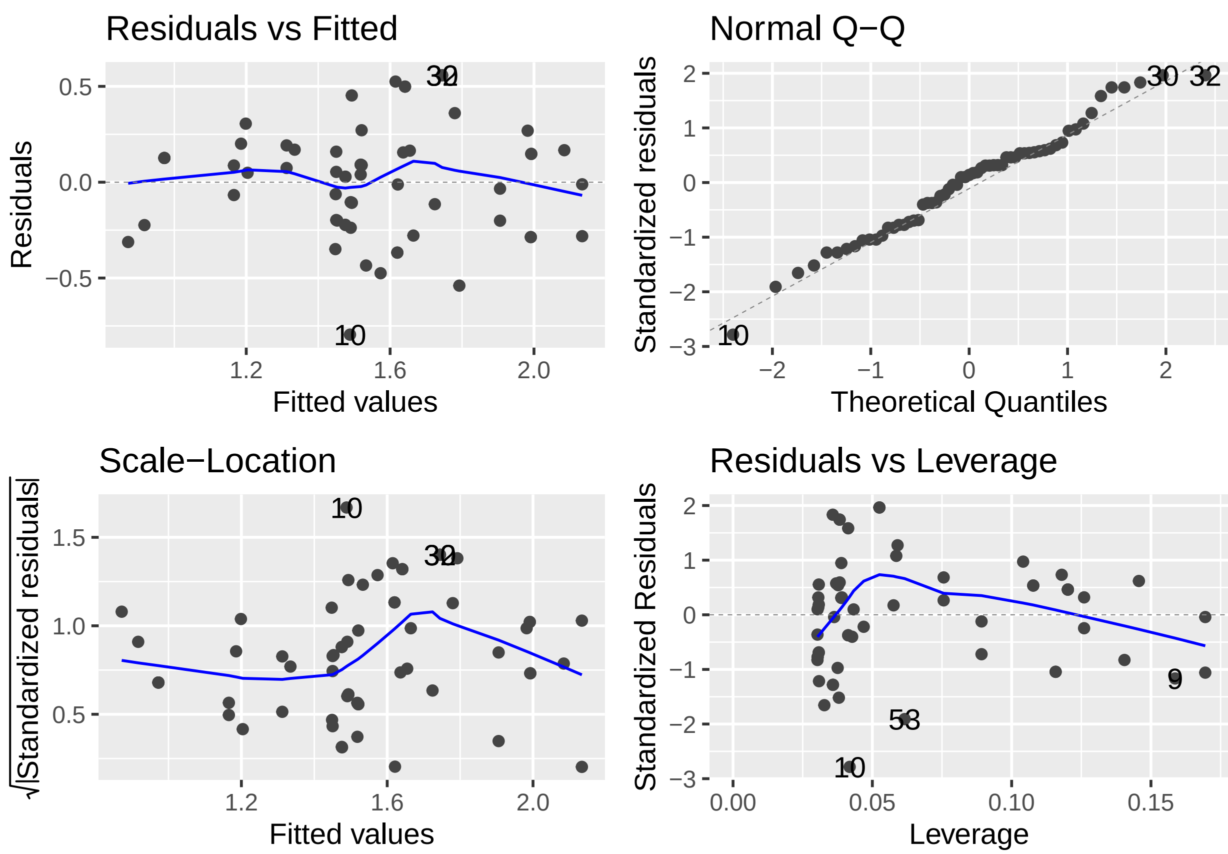

```
autoplot(asian.linear)
```



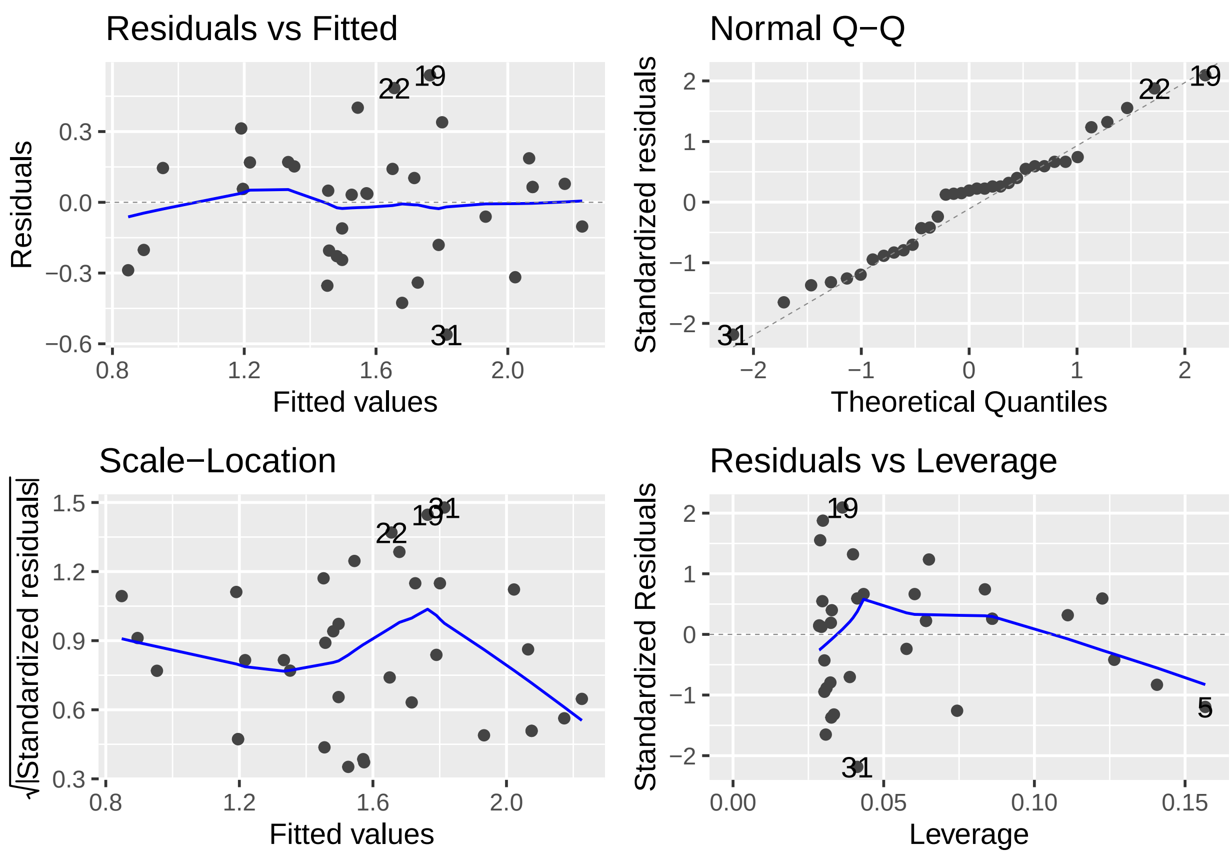

```
autoplot(asian.int)
```



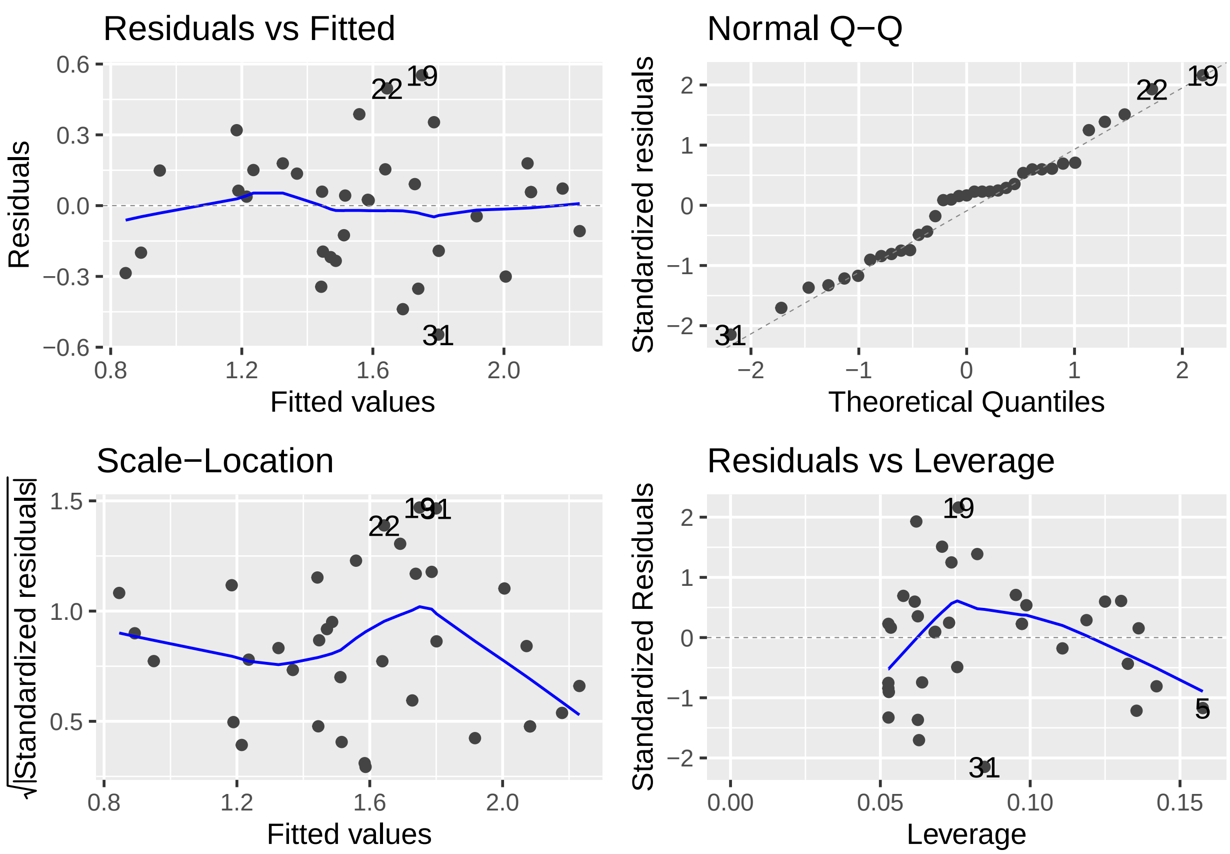

```
autoplot(asian.group)
```



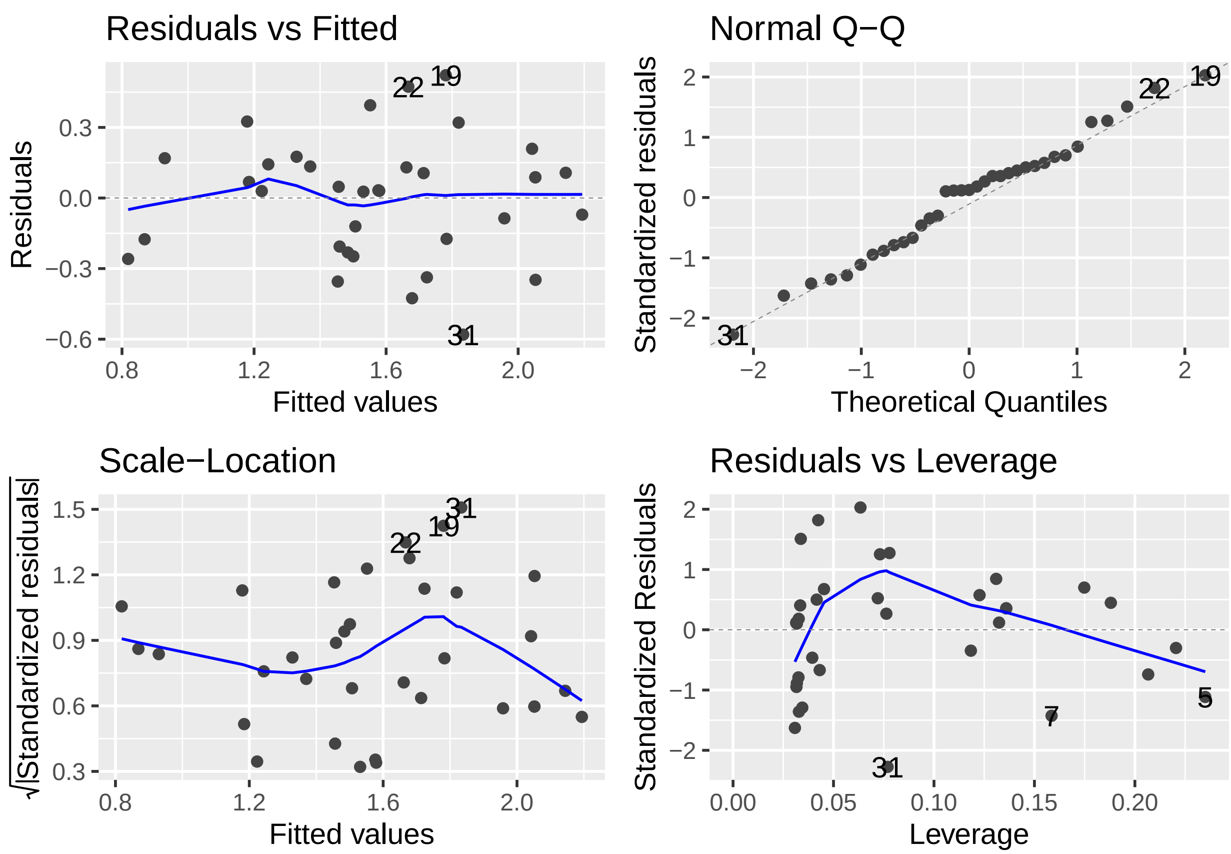

```
# Grass carp
autoplot(grass.simple)
```



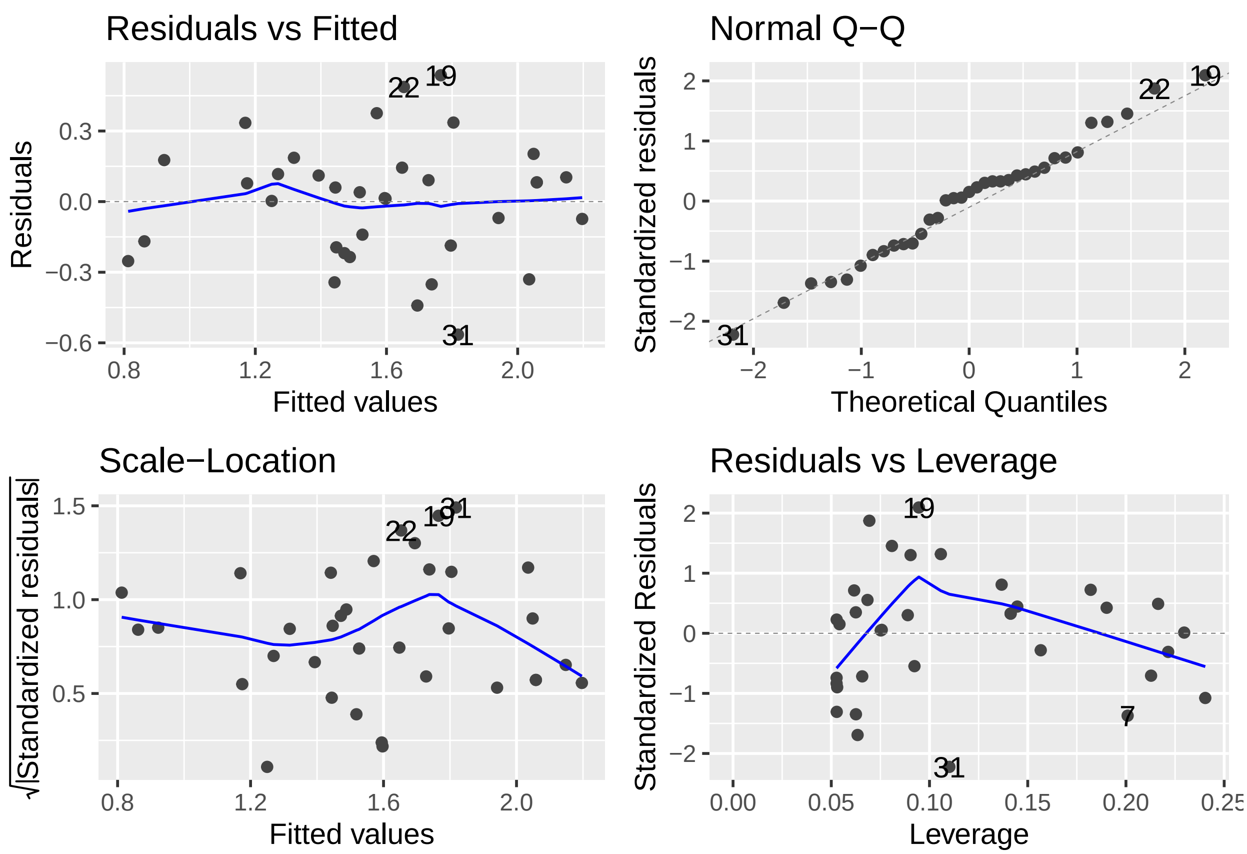

```
autoplot(grass.linear)
```



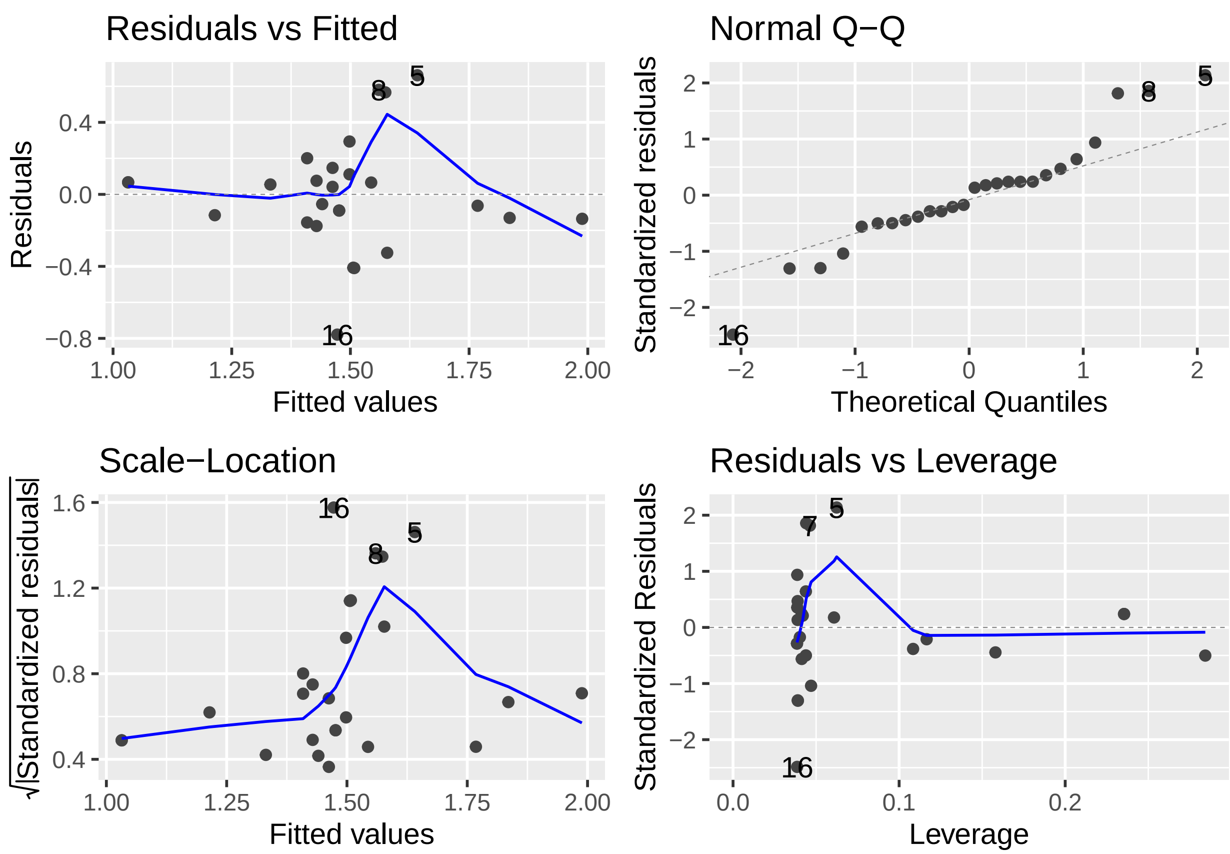

```
autoplot(grass.int)
```



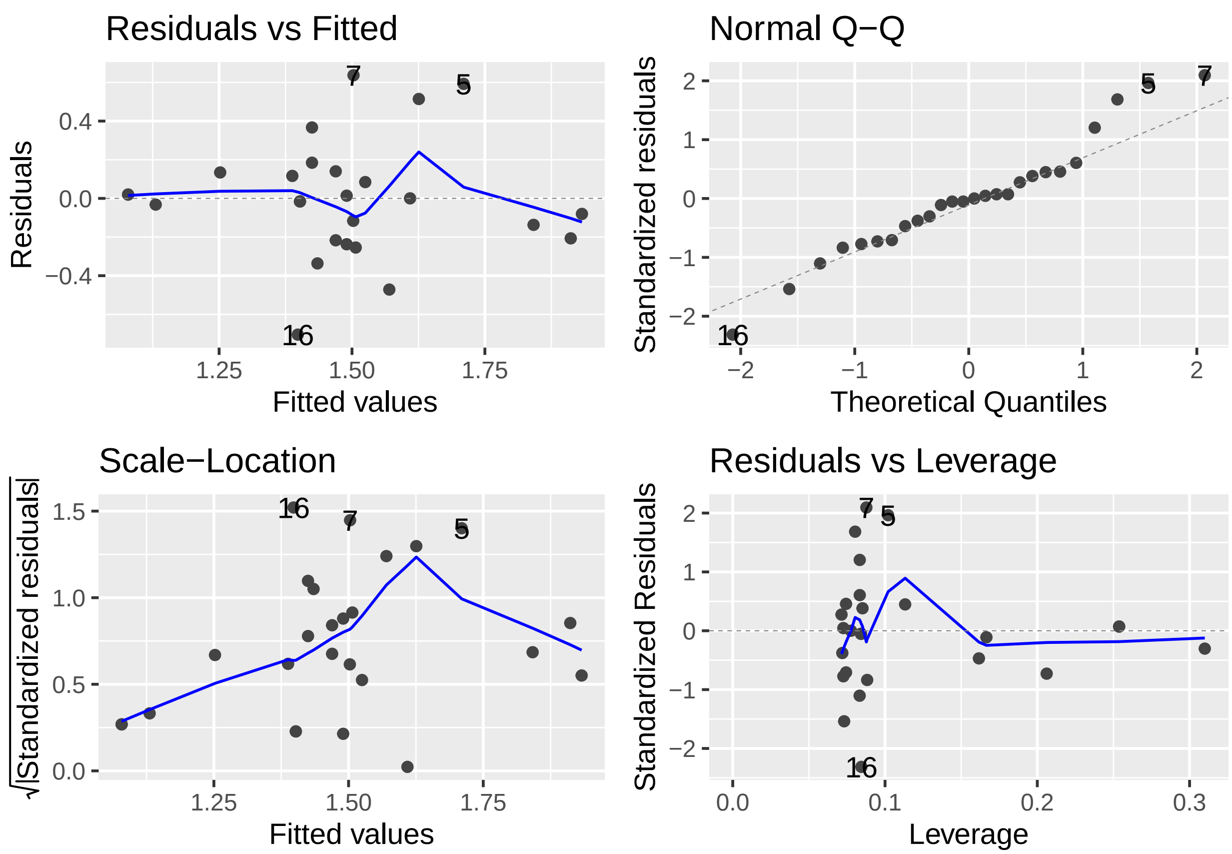

```
autoplot(grass.group)
```



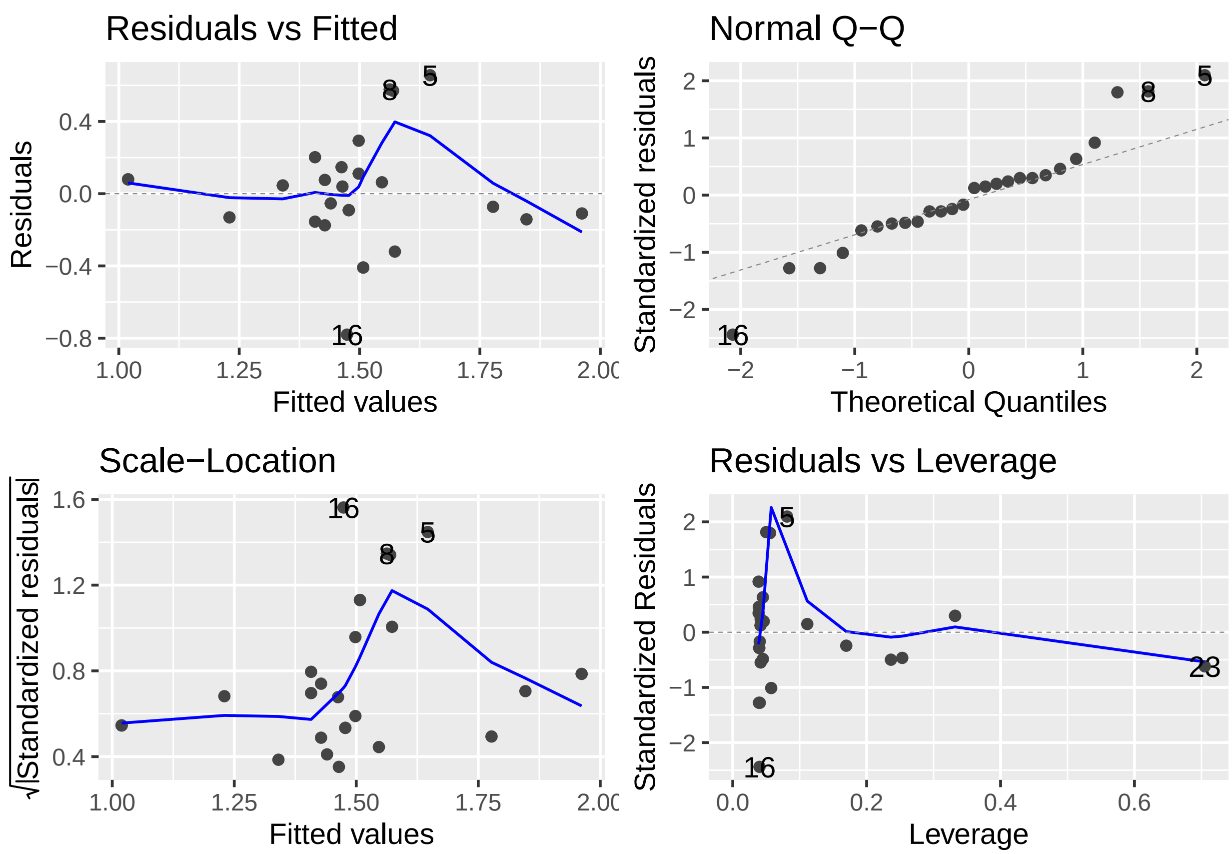

```r
# Bighead and silver carp
autoplot(big.sil.simple)
```



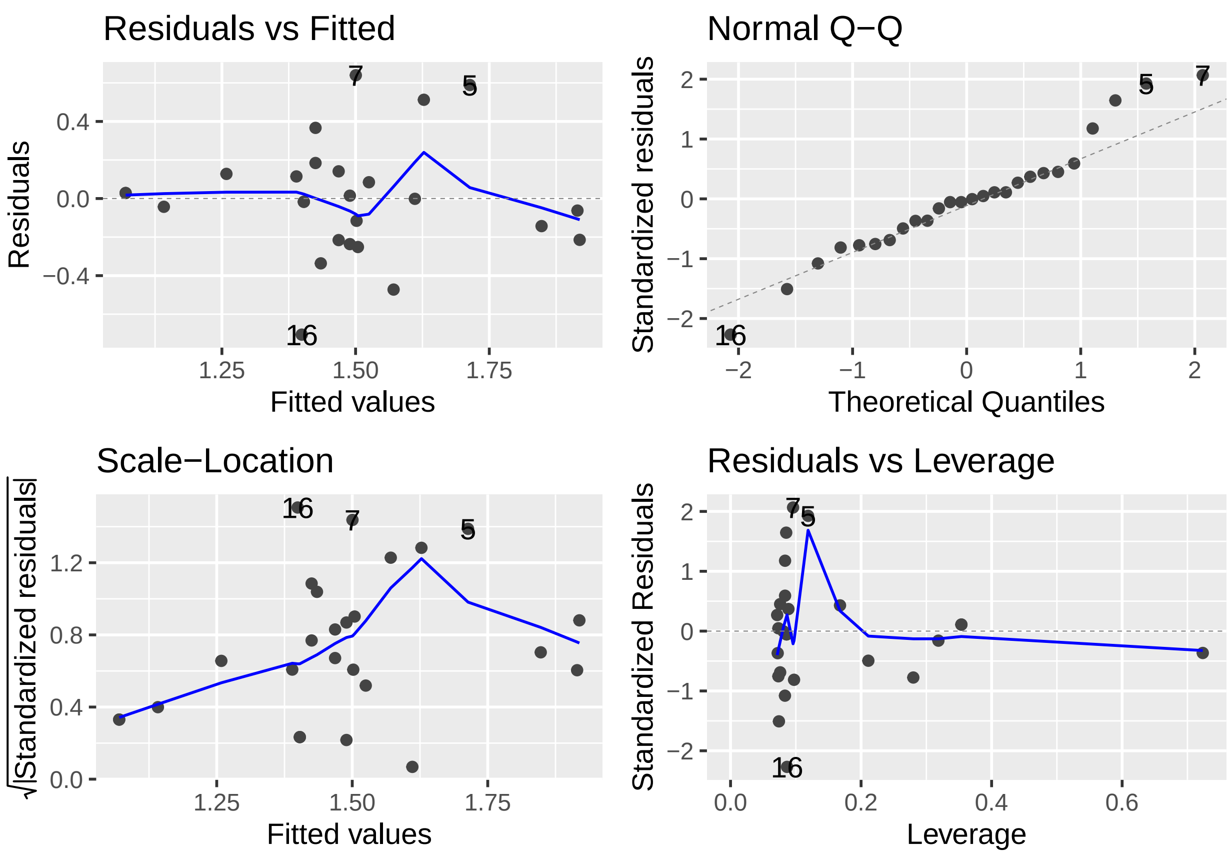

```
autoplot(big.sil.linear)
```



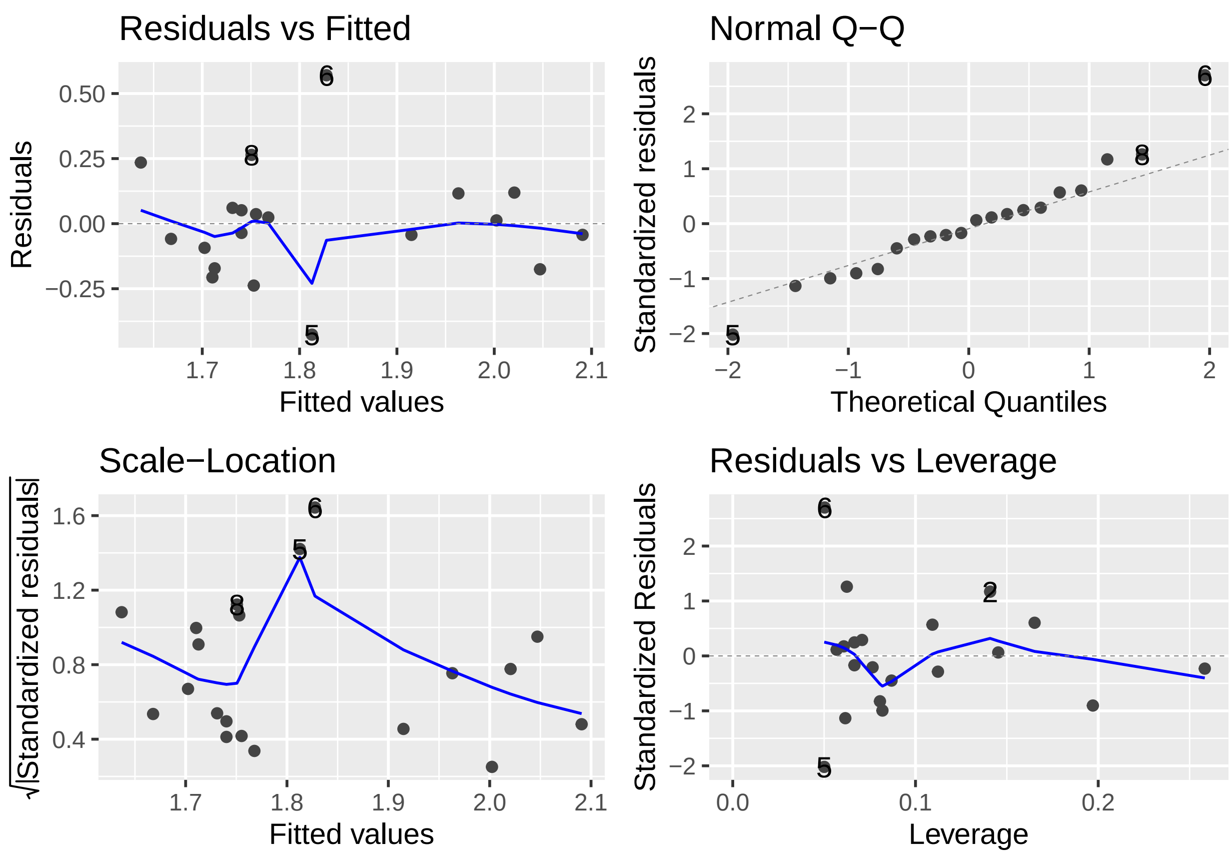

```
autoplot(big.sil.int)
```



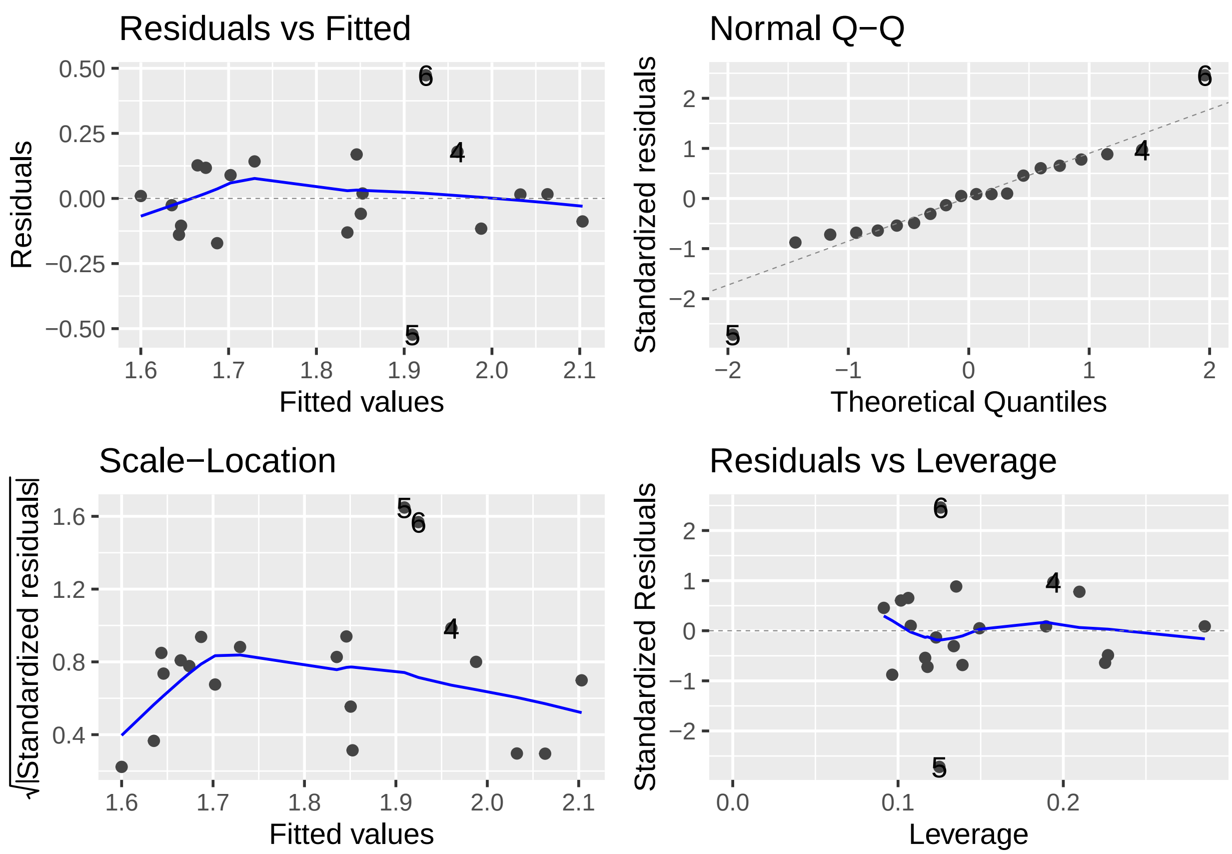

```
autoplot(big.sil.group)
```



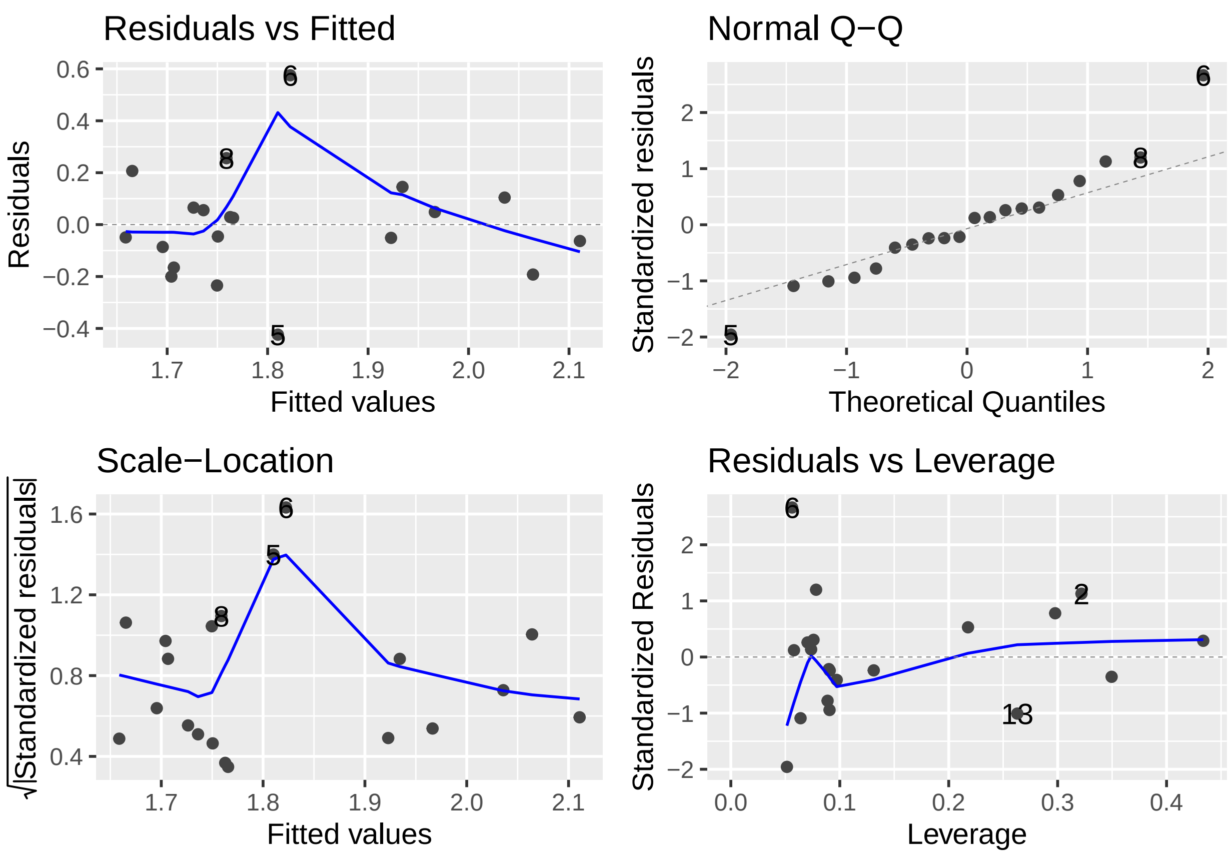

```
# Black carp
autoplot(black.simple)
```



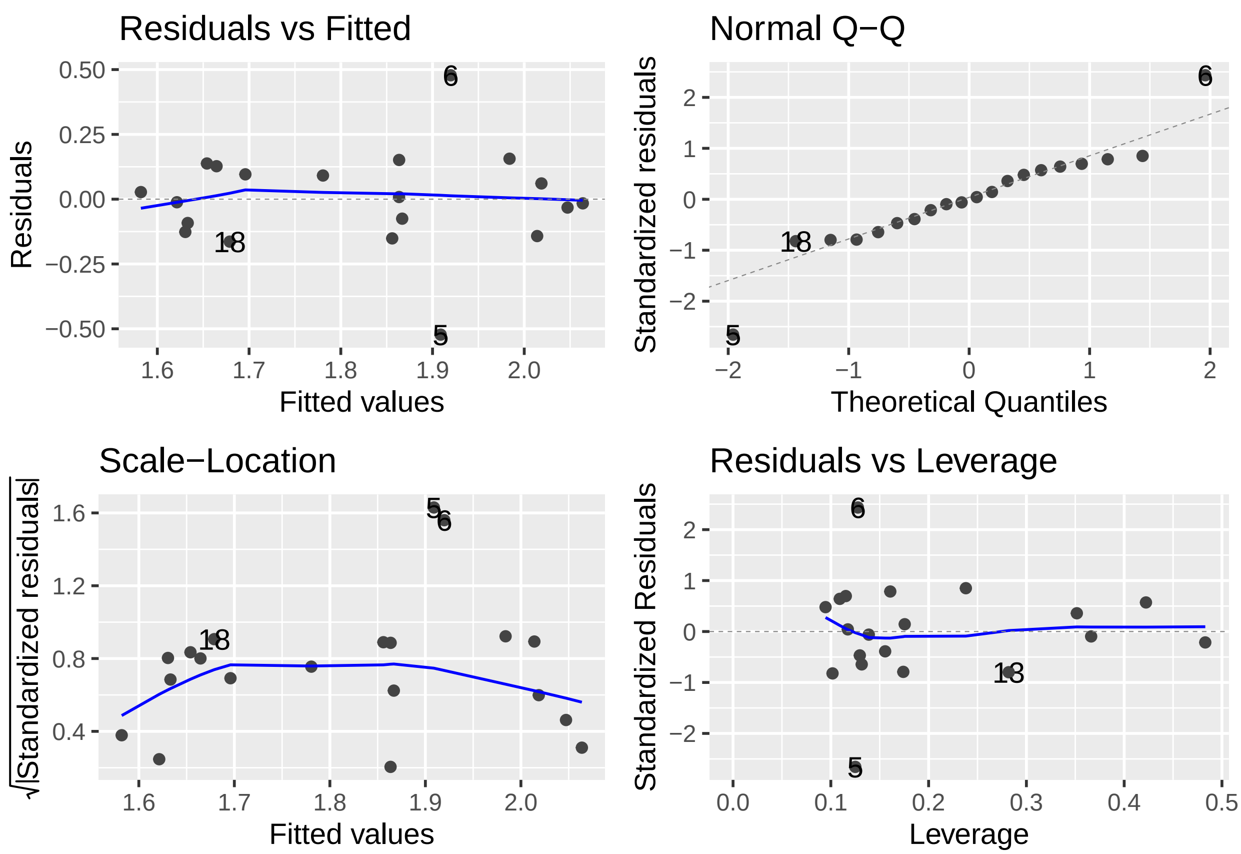

```
autoplot(black.linear)
```



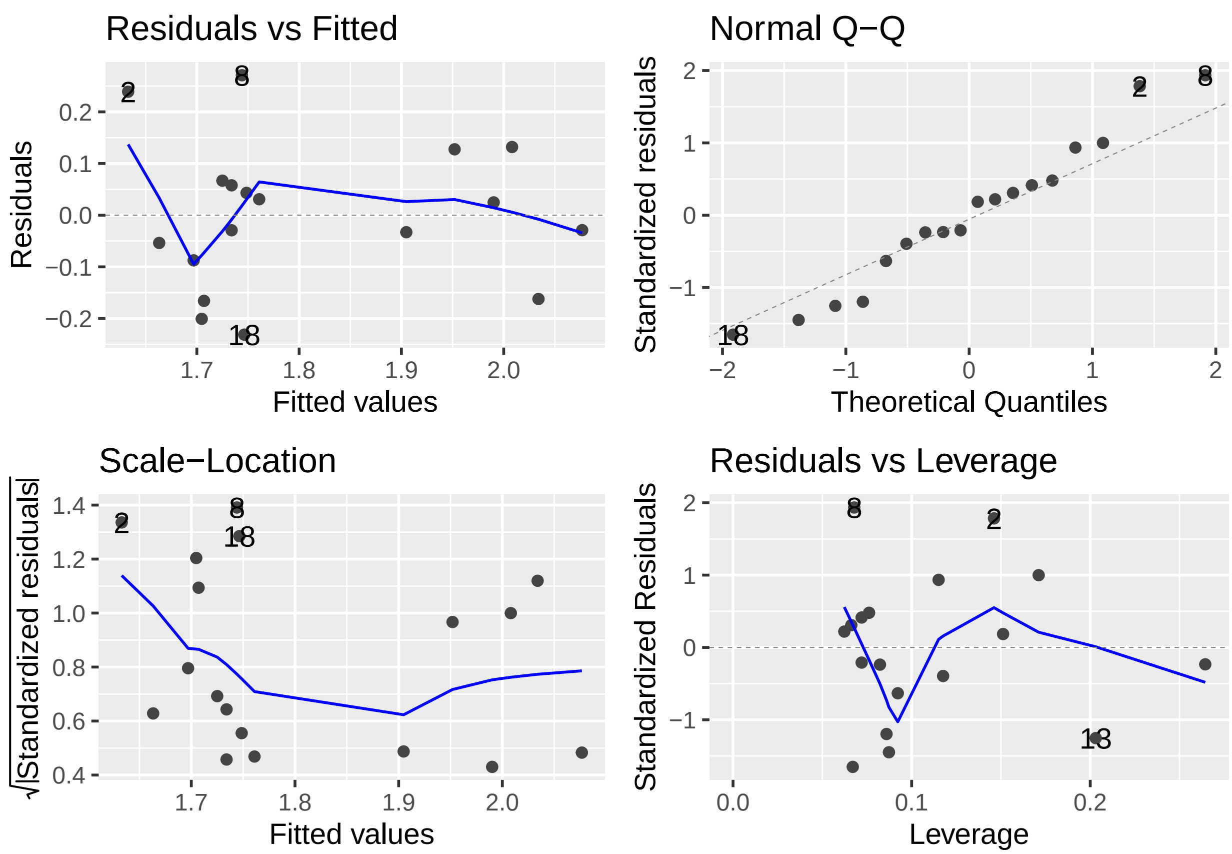

```
autoplot(black.int)
```



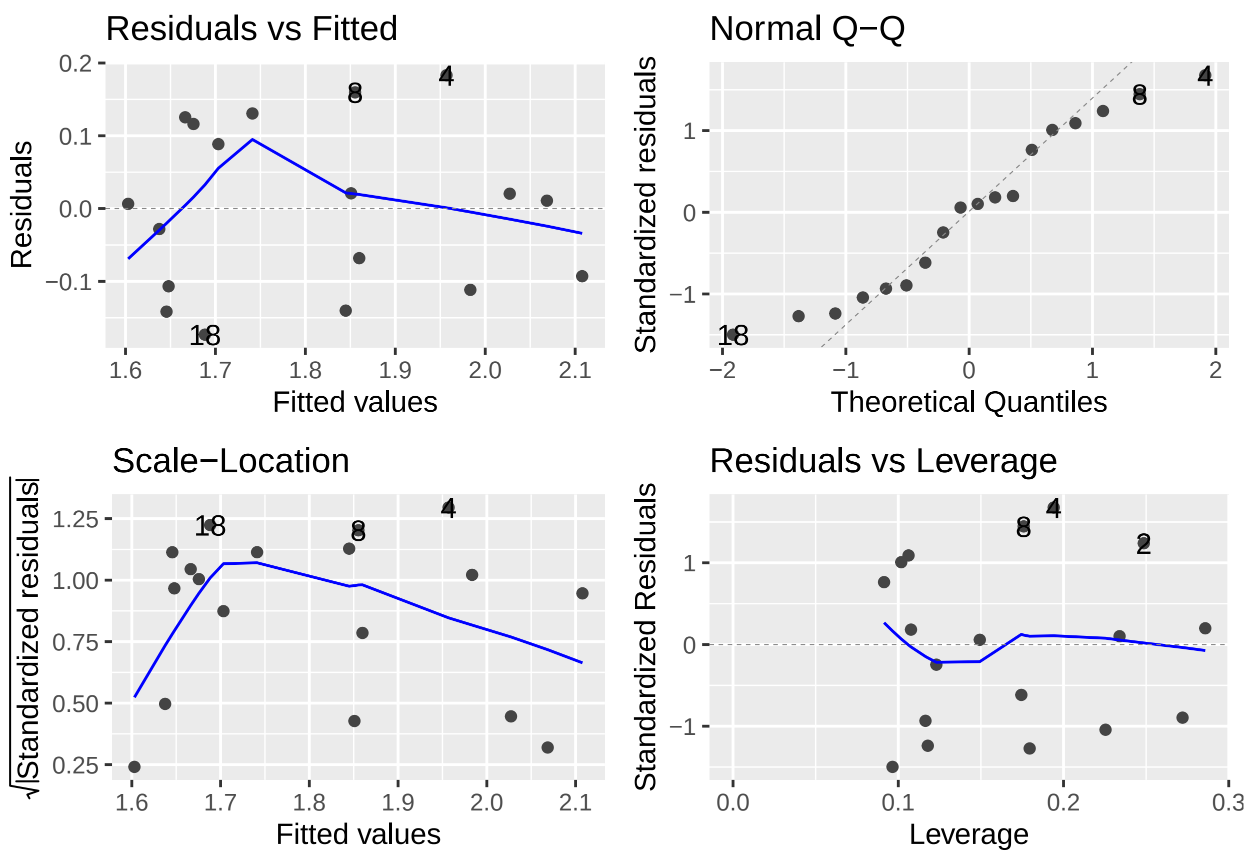

```
autoplot(black.group)
```



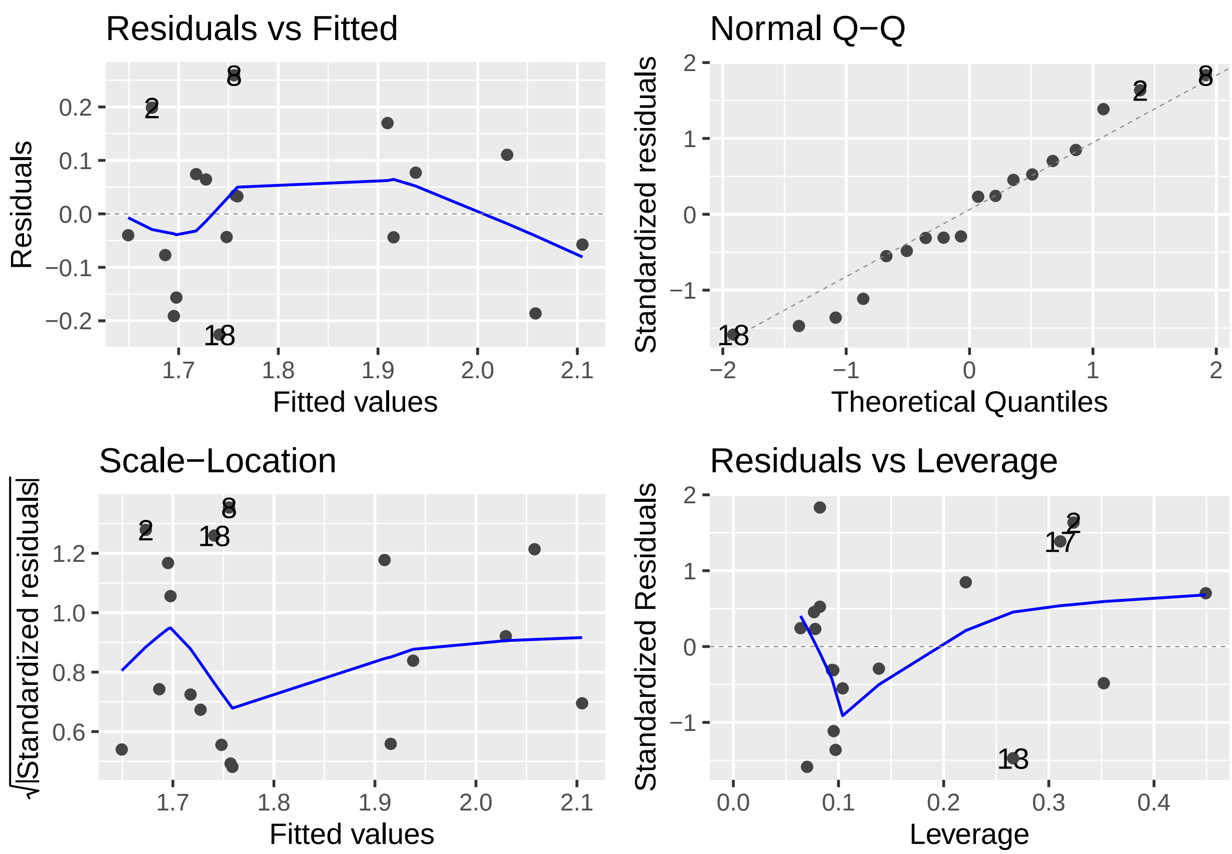

```r
# Black carp modified
autoplot(black.simple.r)
```



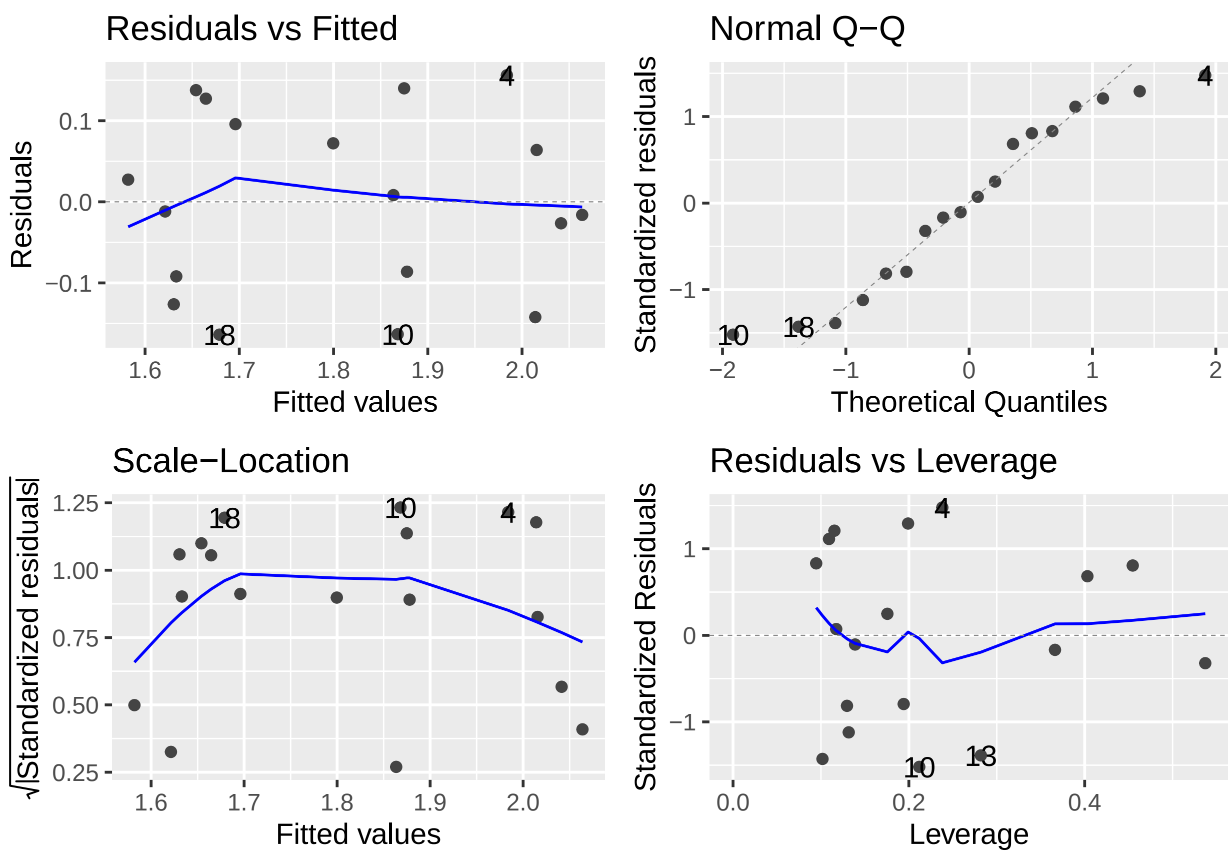

```
autoplot(black.linear.r)
```




```
autoplot(black.int.r)
```




```
autoplot(black.group.r)
```



Residuals vs Fitted
Residuals
0.1
0.0
−0.1
1.6
1.7
1.8
1.9
2.0
Fitted values
4
18
10
Normal Q−Q
Standardized residuals
1
0
−1
−2
−1
0
1
2
Theoretical Quantiles
4
10
18
Scale−Location
√|Standardized residuals|
1.25
1.00
0.75
0.50
0.25
1.6
1.7
1.8
1.9
2.0
Fitted values
18
10
4
Residuals vs Leverage
Standardized Residuals
1
0
−1
0.0
0.2
0.4
Leverage
4
10
18