Teste de erro explicativo

Felipe N. S. Bezerra

20 de outubro de 2018

set.seed(8)

Gerando valores aleatórios para X1

X1 <- rnorm(50,mean=3,sd=10)  
X1

## [1] 2.1541393 11.4040013 -1.6348277 -2.5083500 10.3604043  
## [6] 1.9211860 1.2971085 -7.8833171 -27.1105168 -2.9317433  
## [11] -4.5979380 5.9204986 7.2139859 -9.9448908 3.6928509  
## [16] -5.1303848 18.1085307 0.2839112 18.5825393 0.6265013  
## [21] 15.8312288 2.9051036 -1.0008278 3.2198562 20.4276738  
## [26] -8.0717416 -7.6048948 22.5116132 9.0270674 -17.2060948  
## [31] 18.0667554 12.6371109 -12.5389141 -4.7364238 15.6077796  
## [36] 7.2845578 11.4458354 -3.9915055 2.3513772 7.7098365  
## [41] 6.6350062 -3.1455591 5.2894929 0.3794317 9.1255599  
## [46] -16.5761249 -3.6350747 5.8378752 5.6871796 -3.4010474

Gerando X2 altamente correlacionado com X1

X2 <- 15 -0.9\*X1 +rnorm(50,mean=0,sd=2)  
X2

## [1] 15.2211772 4.0061062 17.6500757 17.0653231 7.6186516 14.4278705  
## [7] 12.2782553 20.8933012 36.4008263 19.1468958 20.2527328 9.8613055  
## [13] 4.9445298 21.7328169 11.1415386 17.4640401 -2.2799811 18.0228759  
## [19] -0.9307313 11.2400020 -2.1834863 14.4656825 17.4664070 11.1479365  
## [25] -3.1968572 23.4625261 16.0498094 -4.0930585 11.6278061 29.1929612  
## [31] -0.7739419 2.0476138 24.0976238 23.3519961 -2.4096689 10.8744643  
## [37] 4.5073011 16.6895660 14.0862915 2.0320927 10.5674636 20.9688872  
## [43] 9.2340228 13.4227351 4.0841321 31.4566882 18.9305055 9.6576210  
## [49] 8.6070638 17.9223772

cor(X1, X2)

## [1] -0.969517

Regressão de X1 explicado por X2

mod.X1.X2 <- lm(X1 ~ X2)  
summary(mod.X1.X2)

##   
## Call:  
## lm(formula = X1 ~ X2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.3305 -1.8560 0.1309 1.6074 5.6091   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 15.75665 0.60410 26.08 <2e-16 \*\*\*  
## X2 -1.06114 0.03871 -27.41 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.525 on 48 degrees of freedom  
## Multiple R-squared: 0.94, Adjusted R-squared: 0.9387   
## F-statistic: 751.5 on 1 and 48 DF, p-value: < 2.2e-16

Erros da regressão de X1

e.X1.X2 <- residuals(mod.X1.X2)  
e.X1.X2

## 1 2 3 4 5 6   
## 2.54928419 -0.10160996 1.33771766 -0.15630880 2.68820818 1.47452175   
## 7 8 9 10 11 12   
## -1.43059754 -1.46925673 -4.24080716 1.62913717 1.13638996 0.62807139   
## 13 14 15 16 17 18   
## -3.29582679 -2.63998703 -0.24107028 -2.35524914 -0.06749673 3.65204961   
## 19 20 21 22 23 24   
## 1.83825420 -3.20293645 -2.24240418 2.49856321 1.77681942 -0.70727586   
## 25 26 27 28 29 30   
## 1.27871284 1.06862519 -6.33045544 2.41165748 5.60914387 -1.98493637   
## 31 32 33 34 35 36   
## 1.48884579 -0.94673421 -2.72462001 4.28665523 -2.70586482 3.06723341   
## 37 38 39 40 41 42   
## 0.47206196 -2.03819503 1.54224996 -5.89047866 2.09191129 3.34870860   
## 43 44 45 46 47 48   
## -0.66856885 -1.13382161 -2.29725490 1.04716396 0.69618540 0.32931024   
## 49 50   
## -0.93617323 -0.13955216

cor(X1,e.X1.X2)

## [1] 0.2450241

cor(X2, e.X1.X2)

## [1] 6.756291e-17

Gerando Y1 como uma função de X1

Y1 <- -9 +1.3\*X1 +rnorm(50,mean=0,sd=4)  
Y1

## [1] -5.0122138 -1.7770927 -17.7147386 -19.3744768 4.6082992  
## [6] -8.4806440 -4.8835792 -20.7057716 -43.5719344 -10.5447757  
## [11] -15.6000416 -4.6552881 2.5662396 -26.2805127 1.6079348  
## [16] -10.1215981 15.5814420 -2.8564117 10.3923359 -9.3791057  
## [21] 7.0029597 -12.1095065 -11.7798183 2.4741152 19.4663352  
## [26] -18.9310700 -25.7904611 21.3712238 -2.7158511 -28.0262280  
## [31] 23.3892750 8.1173120 -31.3029595 -20.5438204 1.6857843  
## [36] 3.6440067 7.6551513 -12.1736528 -0.2308710 -6.3055604  
## [41] 4.9966393 -3.3334495 0.5840852 -16.3826613 3.1847536  
## [46] -33.3680393 -11.3365878 1.7723562 -5.7300108 -15.0994689

Modelo explicando Y1 com X1 e X2

mod.Y1.X1X2 <- lm(Y1 ~ X1 + X2)  
summary(mod.Y1.X1X2)

##   
## Call:  
## lm(formula = Y1 ~ X1 + X2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.6834 -3.1912 0.5708 3.4546 8.9019   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -16.7684 4.4079 -3.804 0.00041 \*\*\*  
## X1 1.7513 0.2704 6.477 5.08e-08 \*\*\*  
## X2 0.4964 0.2959 1.677 0.10012   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.729 on 47 degrees of freedom  
## Multiple R-squared: 0.8936, Adjusted R-squared: 0.8891   
## F-statistic: 197.4 on 2 and 47 DF, p-value: < 2.2e-16

Modelo explicando Y1 com X1 e o erro do primeiro modelo

mod.Y1.X1e1 <- lm(Y1 ~ X1 + e.X1.X2)  
summary(mod.Y1.X1e1)

##   
## Call:  
## lm(formula = Y1 ~ X1 + e.X1.X2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.6834 -3.1912 0.5708 3.4546 8.9019   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -9.39814 0.68858 -13.649 <2e-16 \*\*\*  
## X1 1.28353 0.06833 18.784 <2e-16 \*\*\*  
## e.X1.X2 0.46775 0.27887 1.677 0.1   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.729 on 47 degrees of freedom  
## Multiple R-squared: 0.8936, Adjusted R-squared: 0.8891   
## F-statistic: 197.4 on 2 and 47 DF, p-value: < 2.2e-16

Modelo explicando Y1 com X1 e o erro do primeiro modelo

mod.Y1.X2e1 <- lm(Y1 ~ X2 + e.X1.X2)  
summary(mod.Y1.X2e1)

##   
## Call:  
## lm(formula = Y1 ~ X2 + e.X1.X2)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.6834 -3.1912 0.5708 3.4546 8.9019   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 10.82591 1.13160 9.567 1.31e-12 \*\*\*  
## X2 -1.36200 0.07251 -18.784 < 2e-16 \*\*\*  
## e.X1.X2 1.75128 0.27037 6.477 5.08e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.729 on 47 degrees of freedom  
## Multiple R-squared: 0.8936, Adjusted R-squared: 0.8891   
## F-statistic: 197.4 on 2 and 47 DF, p-value: < 2.2e-16