Contrasting contrasts

Livio Finos

11 / 11 / 2019

Contents


```
knitr::opts_chunk$set(echo = TRUE)
```

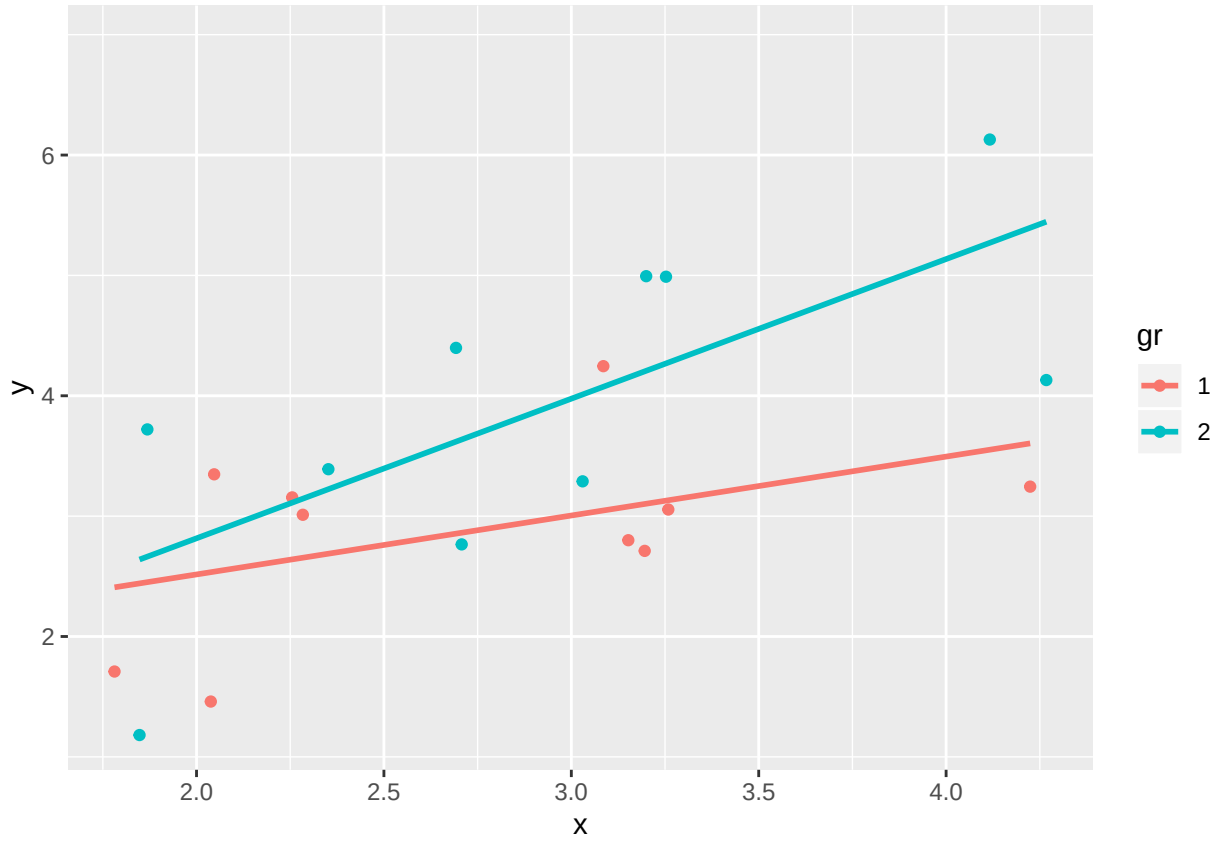
0.1 The data + EDA

```
library(MASS)
set.seed(3)
n0=10
D=data.frame(gr=as.factor(rep(1:2,n0)),
             x=rnorm(n0*2)+3)
D$y=D$x+rnorm(n0*2)
D$y[D$gr=="2"]=D$y[D$gr=="2"]+1

D
```

```
##    gr        x        y
## 1   1 2.038067 1.459583
## 2   2 2.707474 2.765174
## 3   1 3.258788 3.055060
## 4   2 1.847868 1.181393
## 5   1 3.195783 2.711328
## 6   2 3.030124 3.289051
## 7   1 3.085418 4.246034
## 8   2 4.116610 6.128677
## 9   1 1.781143 1.709064
## 10  2 4.267369 4.130586
## 11  1 2.255218 3.155843
## 12  2 1.868781 3.720552
## 13  1 2.283642 3.011357
## 14  2 3.252652 4.989155
## 15  1 3.152046 2.799916
## 16  2 2.692344 4.397859
## 17  1 2.046983 3.347341
## 18  2 2.351757 3.390009
## 19  1 4.224314 3.245030
## 20  2 3.199812 4.993573
```

```
library(ggplot2)
ggplot(D,aes(x=x,y=y,color=gr))+geom_point()+
  geom_smooth(method = "lm", fill = NA)
```



0.2 Un modello lineare

```
modDU=lm(y~gr*x,data=D)
summary(modDU)
```

```
##
## Call:
## lm(formula = y ~ gr * x, data = D)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45800 -0.70502  0.04468  0.78073  1.19905
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.5367     1.1174   1.375    0.188
## gr2          -1.0406     1.5798  -0.659    0.519
## x             0.4895     0.3952   1.239    0.233
## gr2:x         0.6704     0.5399   1.242    0.232
```

```
##
## Residual standard error: 0.9101 on 16 degrees of freedom
## Multiple R-squared:  0.5268, Adjusted R-squared:  0.4381
## F-statistic: 5.937 on 3 and 16 DF,  p-value: 0.006383
```

Le variabili usate nel modello lineare

```
(mm <- model.matrix(~gr*x,data=D))
```

```
##    (Intercept) gr2        x    gr2:x
## 1            1   0 2.038067 0.000000
## 2            1   1 2.707474 2.707474
## 3            1   0 3.258788 0.000000
## 4            1   1 1.847868 1.847868
## 5            1   0 3.195783 0.000000
## 6            1   1 3.030124 3.030124
## 7            1   0 3.085418 0.000000
## 8            1   1 4.116610 4.116610
## 9            1   0 1.781143 0.000000
## 10           1   1 4.267369 4.267369
## 11           1   0 2.255218 0.000000
## 12           1   1 1.868781 1.868781
## 13           1   0 2.283642 0.000000
## 14           1   1 3.252652 3.252652
## 15           1   0 3.152046 0.000000
## 16           1   1 2.692344 2.692344
## 17           1   0 2.046983 0.000000
## 18           1   1 2.351757 2.351757
## 19           1   0 4.224314 0.000000
## 20           1   1 3.199812 3.199812
## attr(,"assign")
## [1] 0 1 2 3
## attr(,"contrasts")
## attr(,"contrasts")$gr
## [1] "contr.treatment"
```

Notate il Multiple R-squared delle prime tre colonne per spiegare la colonna dell'interazione:

```
summary(lm(mm[,4]~mm[,-4]+0))
```

```
##
## Call:
## lm(formula = mm[, 4] ~ mm[, -4] + 0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79942 -0.25381 -0.03003  0.28340  0.61927
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## mm[, -4](Intercept)  -1.4637     0.3549  -4.125 0.000708 ***
## mm[, -4]gr2           2.8256     0.1845  15.318 2.22e-11 ***
```

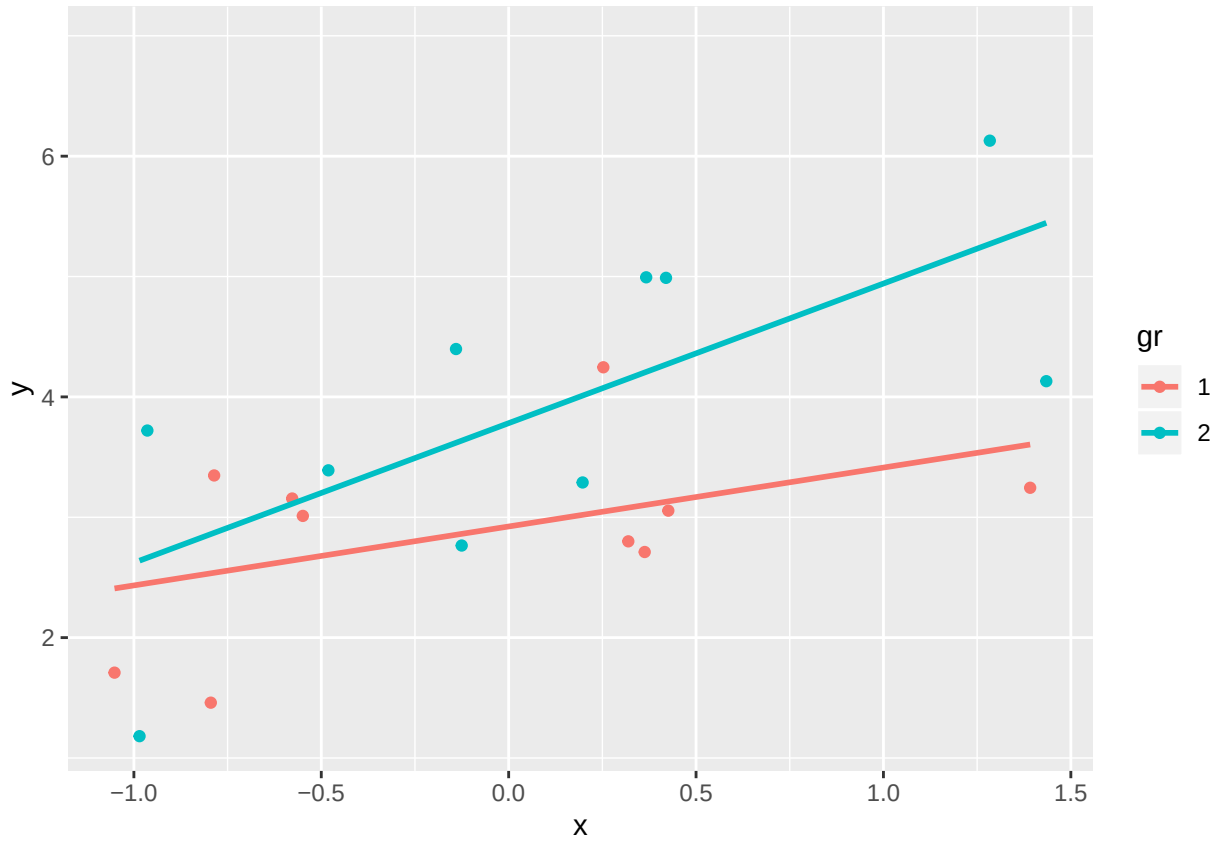
```
## mm[, -4]x             0.5357     0.1210   4.429 0.000367 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4089 on 17 degrees of freedom
## Multiple R-squared:  0.9692, Adjusted R-squared:  0.9637
## F-statistic: 178.1 on 3 and 17 DF,  p-value: 4.85e-13
```

0.3 Un secondo modello lineare

```
D2=D
D2$x=D$x-mean(D$x)
modDUC=lm(y~gr*x,data=D2)
summary(modDUC)
```

```
##
## Call:
## lm(formula = y ~ gr * x, data = D2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45800 -0.70502  0.04468  0.78073  1.19905
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.9233     0.2905  10.062 2.52e-08 ***
## gr2           0.8585     0.4106   2.091   0.0529 .
## x             0.4895     0.3952   1.239   0.2333
## gr2:x         0.6704     0.5399   1.242   0.2322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9101 on 16 degrees of freedom
## Multiple R-squared:  0.5268, Adjusted R-squared:  0.4381
## F-statistic: 5.937 on 3 and 16 DF,  p-value: 0.006383
```

```
ggplot(D2,aes(x=x,y=y,color=gr))+geom_point()+
  geom_smooth(method = "lm", fill = NA)
```



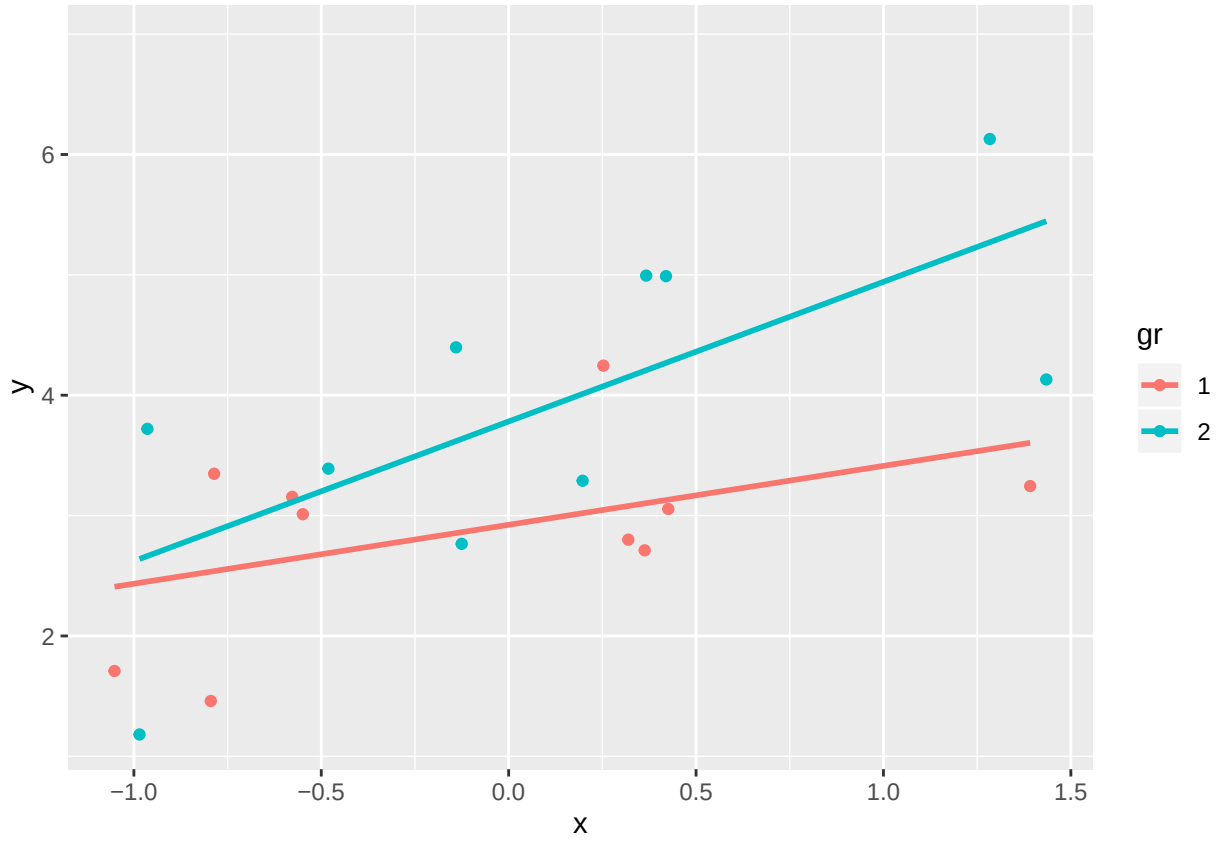
1 . . . e un terzo

```
contrasts(D2$gr)=contr.sum(2)
modS0=lm(y~gr*x,data=D2)
summary(modS0)
```

```
##
## Call:
## lm(formula = y ~ gr * x, data = D2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.45800 -0.70502  0.04468  0.78073  1.19905
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.3526     0.2053  16.329 2.13e-11 ***
## gr1          -0.4293     0.2053  -2.091  0.05287 .
## x             0.8247     0.2699   3.055  0.00756 **
## gr1:x        -0.3352     0.2699  -1.242  0.23223
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.9101 on 16 degrees of freedom
## Multiple R-squared:  0.5268, Adjusted R-squared:  0.4381
## F-statistic: 5.937 on 3 and 16 DF,  p-value: 0.006383
```

```
ggplot(D2,aes(x=x,y=y,color=gr))+geom_point()+
  geom_smooth(method = "lm", fill = NA)
```




```
(mm <- model.matrix(~gr*x,data=D2))
```

```
##    (Intercept) gr1          x      gr1:x
## 1            1   1 -0.7947430 -0.7947430
## 2            1  -1 -0.1253353  0.1253353
## 3            1   1  0.4259787  0.4259787
## 4            1  -1 -0.9849415  0.9849415
## 5            1   1  0.3629733  0.3629733
## 6            1  -1  0.1973144 -0.1973144
## 7            1   1  0.2526082  0.2526082
## 8            1  -1  1.2838006 -1.2838006
## 9            1   1 -1.0516670 -1.0516670
## 10           1  -1  1.4345592 -1.4345592
## 11           1   1 -0.5775912 -0.5775912
## 12           1  -1 -0.9640281  0.9640281
## 13           1   1 -0.5491681 -0.5491681
## 14           1  -1  0.4198428 -0.4198428
## 15           1   1  0.3192361  0.3192361
```

```
## 16           1  -1 -0.1404660  0.1404660
## 17           1   1 -0.7858269 -0.7858269
## 18           1  -1 -0.4810524  0.4810524
## 19           1   1  1.3915041  1.3915041
## 20           1  -1  0.3670020 -0.3670020
## attr(,"assign")
## [1] 0 1 2 3
## attr(,"contrasts")
## attr(,"contrasts")$gr
##   [,1]
## 1    1
## 2   -1
```

Notate il Multiple R-squared delle prime tre colonne per spiegare la colonna dell'interazione:

```
summary(lm(mm[,4]~mm[,-4]+0))
```

```
##
## Call:
## lm(formula = mm[, 4] ~ mm[, -4] + 0)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.23853 -0.56680  0.06006  0.50762  1.59884
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## mm[, -4](Intercept) -0.100670   0.182843  -0.551    0.589
## mm[, -4]gr1         -0.007197   0.184458  -0.039    0.969
## mm[, -4]x           -0.071487   0.241915  -0.296    0.771
##
## Residual standard error: 0.8177 on 17 degrees of freedom
## Multiple R-squared:  0.02245,    Adjusted R-squared:  -0.1501
## F-statistic: 0.1302 on 3 and 17 DF,  p-value: 0.9408
```