



CONDOMINI

Contrasti pt.2

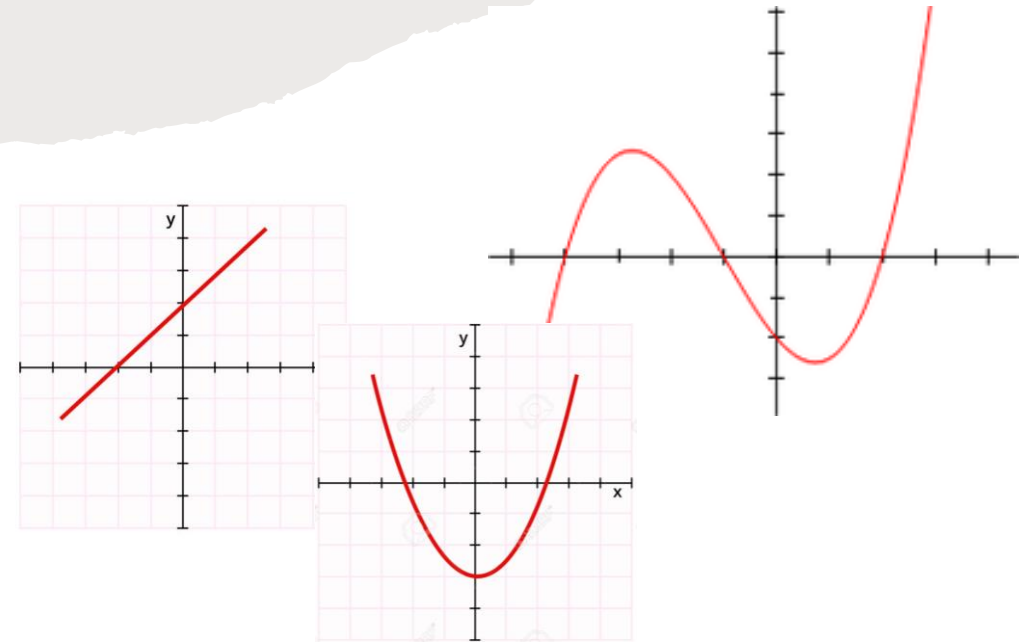
Gruppo 5:
Francesco Costalunga
Chiara Borsi
Martino Spelzini
Jacopo Colangelo

Contrasti polinomiali

Rilevare un andamento di tipo lineare, quadratico, cubico nella nostra distribuzione di dati, spesso rilevante in studi longitudinali

La matrice dei contrasti, lavorerà in ordine crescente (n^2, n^3, \dots), continuando ad aumentare in base al numero dei contrasti ($n-1$). È normalizzata e corrispondente alla trasposta della matrice delle ipotesi

Ogni confronto farà riferimento alla media globale, e sarà di tipo ortogonale

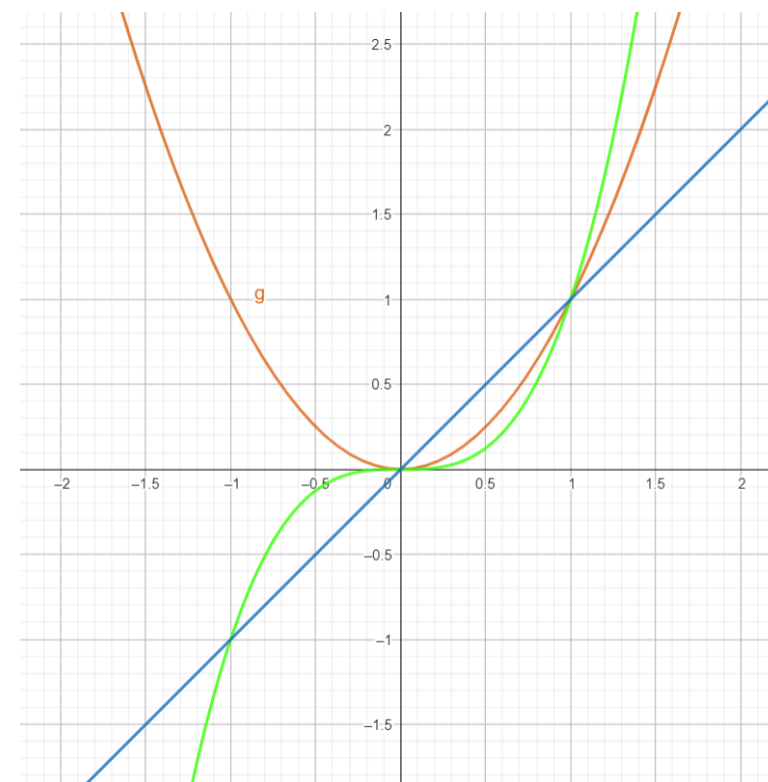


```
C      .L      .Q
[1,] -7.071068e-01  0.4082483
[2,] -7.850462e-17 -0.8164966
[3,]  7.071068e-01  0.4082483
> t(contr.poly(3))
      [,1]      [,2]      [,3]
.L -0.7071068 -7.850462e-17  0.7071068
.Q  0.4082483 -8.164966e-01  0.4082483
```

Contrasti polinomiali su Rstudio

Funzione `contr.poly(n)`

	.L	.Q	.C
[1,]	-0.6708204	0.5	-0.2236068
[2,]	-0.2236068	-0.5	0.6708204
[3,]	0.2236068	-0.5	-0.6708204
[4,]	0.6708204	0.5	0.2236068



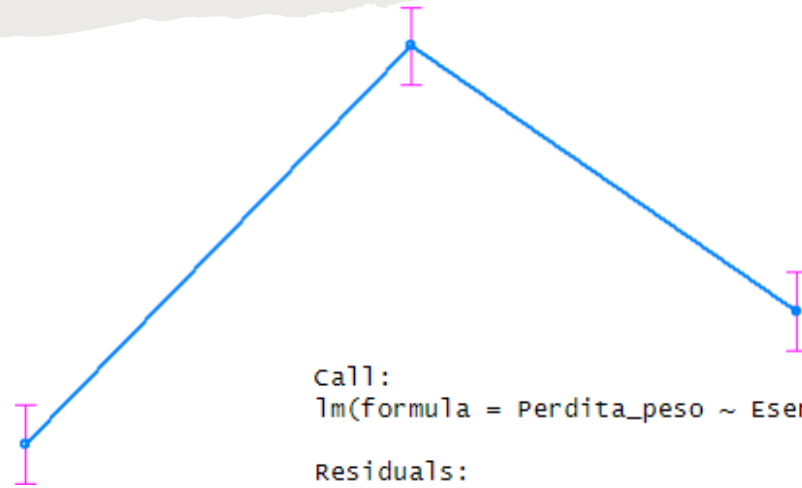
Output Rstudio

Summary

- Effetto lineare
- Effetto quadratico
- ...

Effect plot

Funzione a livello grafico



```
Call:
lm(formula = Perdita_peso ~ Esercizio, data = esercizio_pp)

Residuals:
    Min       1Q   Median       3Q      Max
-0.3800 -0.1625 -0.0100  0.1300  0.4200

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.09667    0.03914   28.018  < 2e-16 ***
Esercizio.L   0.33234    0.06780    4.902 3.96e-05 ***
Esercizio.Q  -0.95938    0.06780  -14.151 5.23e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2144 on 27 degrees of freedom
Multiple R-squared:  0.8926,    Adjusted R-squared:  0.8846
F-statistic: 112.1 on 2 and 27 DF,  p-value: 8.341e-14
```

Contrasti Helmert



Obiettivo principale: Codificare la differenza tra un livello e quelli immediatamente precedenti.



Esempio: Se consideriamo una variabile categoriale a tre livelli A,B,C, il primo contrasto codificherà la differenza tra i primi due livelli (A e B) mentre il secondo codificherà la differenza tra il terzo livello (C) e la media dei primi due (A+B).

Questo può essere dedotto anche dalla matrice dei contrasti.



Qui la prima colonna rappresenta il primo contrasto A VS B dove il livello A = (-1) mentre B=1. Il livello C viene codificato come 0 perché in questo momento non deve essere considerato.

Nella seconda colonna viene rappresentato il secondo contrasto dove A e B = (-1) perché considerati come unico termine di paragone mentre C= 2.

La matrice delle ipotesi codifica in modo visivo le nostre ipotesi di partenza e da essa si ottiene la matrice di contrasto

$$H = \begin{bmatrix} -1/2 & 1/2 & 0 \\ -1/6 & -1/6 & 1/3 \end{bmatrix}$$



Primo contrasto

Secondo contrasto

$$C = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ 0 & 2 \end{bmatrix}$$



$$-1 + 1 = 0$$

$$-1 + (-1) + 2 = 0$$

Applicazione pratica in Rstudio



La funzione associata ai contrasti Helmert fa parte della famiglia dei contrasti (contrasts) ed è: `contr.Helmert`

Passaggi preliminari:

- Dato un database prestabilito (anorexia) trasformare la VI in fattore

```
anorexia$Treat=factor(anorexia$Treat)
```

- Usando la funzione contrasts verificare la matrice dei contrasti

```
contrasts(anorexia$Treat)
```

- ed eventualmente riorganizzare i livelli di interesse

```
anorexia$Treat=factor(anorexia$Treat, levels=c("Cont", "CBT", "FT"))
```

Output in R

→

	Cont	FT
CBT	0	0
Cont	1	0
FT	0	1

→

	CBT	FT
Cont	0	0
CBT	1	0
FT	0	1

- Associare alla matrice di default (treatment) la matrice helmert indicando il numero dei livelli della VI tra parentesi

```
contrasts(anorexia$Treat)=contr.helmert(3)
```

- Ricontrollare la matrice associata

```
contrasts(anorexia$Treat)
```



	[,1]	[,2]
Cont	-1	-1
CBT	1	-1
FT	0	2

- Creare il modello di regressione

```
modello=lm(Postwt~Treat,anorexia)
```

Residuals:

Min	1Q	Median	3Q	Max
-15.2941	-3.7299	-0.0021	4.7809	17.9034

- Verificare i parametri associati al modello e alla sua bontà

```
summary(modello)
```



Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	85.7661	0.8819	97.256	< 2e-16	***
Treat1	2.2944	0.9842	2.331	0.022667	*
Treat2	2.3640	0.6744	3.505	0.000806	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.288 on 69 degrees of freedom

Multiple R-squared: 0.2005, Adjusted R-squared: 0.1773

F-statistic: 8.651 on 2 and 69 DF, p-value: 0.0004443

Contrasti Helmert Invertiti

- Utilizzati per esaminare le **differenze gerarchiche** tra gruppi in un set di dati
- **Strumento specifico** per analisi tra gruppi con dati ordinati
- Esempio **scenario di utilizzo** : confronto trattamento terapia
- **Vantaggi e Applicabilità**

Implementazione in R con Esempio Pratico



1) `anorexia$Treat=factor(anorexia$Treat)`

`contrasts(anorexia$Treat)`

`anorexia$Treat=factor(anorexia$Treat, levels = c("Cont", "CBT", "FT"))`

`contrasts(anorexia$Treat)`

`matrice=matrix(c(2,-1,-1,
0,1,-1),3,2)`

2) `contrasts(anorexia$Treat)=matrice`

`contrasts(anorexia$Treat)`

`modello=lm(Postwt ~ Treat, data=anorexia)`

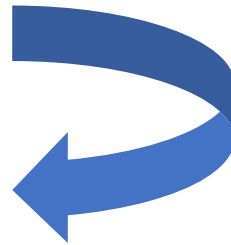
`summary(modello)`

`library(effects)`

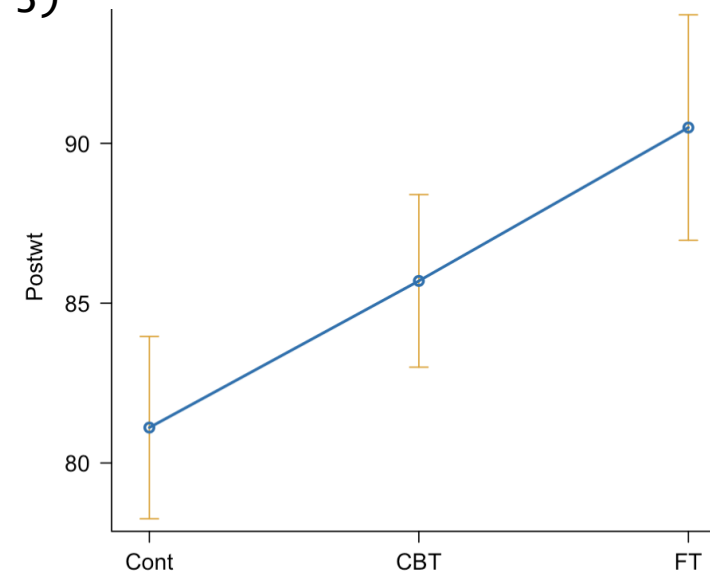
`plot(effect("Treat", modello))`



La funzione specifica per Helmert invertiti non esiste in R



3)



Contrasti Personalizzati

Le **ipotesi** sono l'aspetto fondamentale su cui si basa la scelta dei contrasti

In specifici disegni si potrebbe dover ricorrere a dei contrasti non presenti nelle altre tipologie

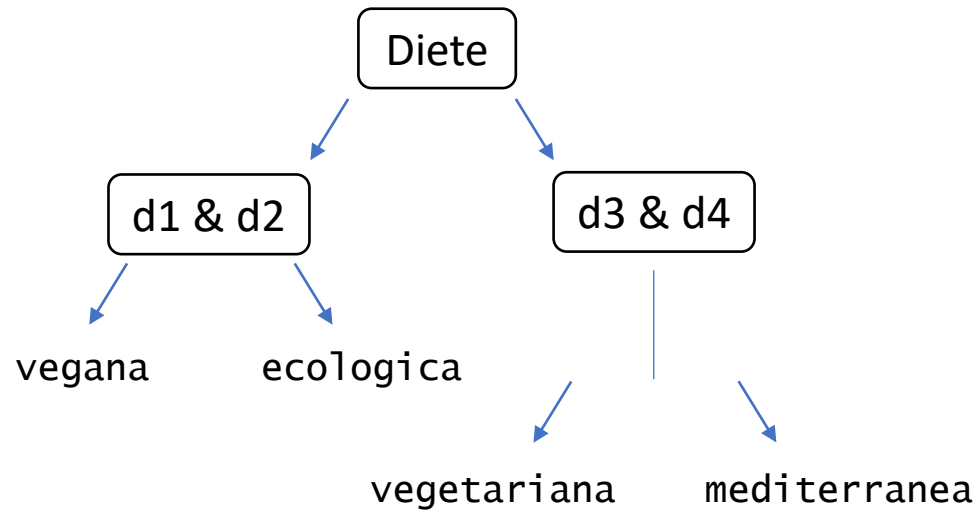
Nella costruzione di una matrice di contrasti personalizzati bisogna mantenere la **ortogonalità**.

- centrati
- interpretazioni singole
- incorrelati



Esempio Pratico

4 livelli di diete da confrontare come impatto di co2. Queste divise in 2 gruppi di impatto principali



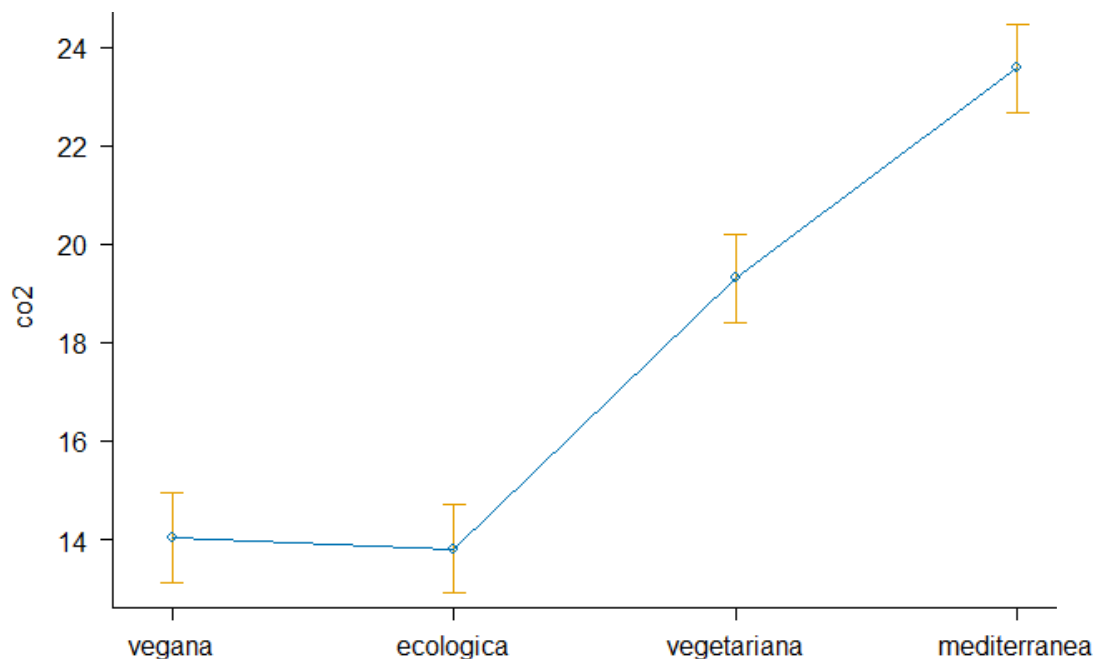
Matrice delle ipotesi

vegana	ecologica	vegetariana	mediterranea
1/2	1/2	-1/2	-1/2
1	-1	0	0
0	0	1	-1

Esempio su Rstudio

Trasformazione da matrice delle ipotesi a matrice di contrasto

`round(ginv(H), 2)`



```
Call:
lm(formula = co2 ~ dieta, data = datax)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-26.0514	-2.6031	0.3036	2.7643	15.3962

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	17.6814	0.2300	76.880	< 2e-16 ***
dieta1	-7.5142	0.4600	-16.336	< 2e-16 ***
dieta2	0.2254	0.6505	0.347	0.729
dieta3	-4.2666	0.6505	-6.559	1.54e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.824 on 436 degrees of freedom
Multiple R-squared: 0.4156, Adjusted R-squared: 0.4115
F-statistic: 103.3 on 3 and 436 DF, p-value: < 2.2e-16

Grazie per l'attenzione!