

Aplicaciones del análisis multivariante con R

Estadística Multivariante - Universidad de Granada

Laura Gómez Garrido

Miguel Lentisco Ballesteros

Antonio Martín Ruiz

Daniel Pozo Escalona

Francisco Javier Sáez Maldonado

22 de enero de 2020

- 1 Introducción: R
- 2 R en el análisis multivariante
 - Distribución Normal Multivariante. Ejemplos
- 3 Para ampliar

1 Introducción: R

2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos

3 Para ampliar

¿Qué es R?

R es un entorno y lenguaje de programación enfocados a la computación estadística y de gráficos. Surge como una reimplementación libre del lenguaje y entorno S. Proporciona una amplia variedad de funcionalidades estadísticas y gráficas y es altamente extensible.

R está disponible como software libre bajo los términos de la GNU General Public License de la Free Software Foundation en forma de código fuente. Puede ser compilado y ejecutado en una gran cantidad de plataformas UNIX, Windows y MacOS.

Entornos de desarrollo para R

Principales librerías

Cómo instalar dependencias

R incluye un gestor de paquetes integrado en el lenguaje, que permite instalar dependencias con la orden `install.packages()`.

```
install.packages("MVA")  
install.packages("HSAUR2")  
install.packages("car")  
install.packages("MASS")
```


Cómo instalar dependencias

Para importar los paquetes para usarlos en un programa:

```
library("MVA")  
library("HSAUR2")  
library("car")  
library("MASS")
```

1 Introducción: R

2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos

3 Para ampliar

Nuestro dataset

Característica del data set	Multivariante	Nº de Instancias	178
Características de los atributos	Enteros, Reales	Nº de Atributos	13
Área	Física	Donado	01/07/1991

Fuente: Machine Learning Repository

Propietarios Originales:

Forina, M. et al, PARVUS -

*An Extendible Package for Data Exploration, Classification
and Correlation.*

*Institute of Pharmaceutical and Food Analysis and Technologies,
Via Brigata Salerno, 16147 Genoa, Italy.*



```
wine <- read.table("http://archive.ics.uci.edu/ml/mach  
sapply(wine[2:14], mean)  
"Media muestral del conjunto completo"
```

In [4]: wine

Attribute (C1-C14)													
V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
4.90	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
5.16	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
5.42	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
5.68	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
5.94	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
6.20	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
6.46	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
6.72	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
6.98	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
7.24	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
7.50	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
7.76	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
8.02	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
8.28	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
8.54	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
8.80	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
9.06	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
9.32	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
9.58	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
9.84	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
10.10	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
10.36	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
10.62	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
10.88	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
11.14	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
11.40	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
11.66	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
11.92	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
12.18	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
12.44	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
12.70	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
12.96	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
13.22	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
13.48	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
13.74	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
14.00	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
14.26	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
14.52	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
14.78	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
15.04	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
15.30	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
15.56	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
15.82	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
16.08	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
16.34	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
16.60	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
16.86	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
17.12	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
17.38	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
17.64	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
17.90	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
18.16	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
18.42	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
18.68	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
18.94	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
19.20	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
19.46	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
19.72	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36
20.00	3.63	3.61	3.41	3.57	3.49	3.37	3.36	3.36	3.36	3.36	3.36	3.36	3.36

Dibujado de datos seleccionados. Scatterplots

```
selection1 <- wine[wine$V1 == "1",]  
selection3 <- wine[wine$V1 == "3",]  
mean1 <- sapply(selection1[2:14], mean)  
mean2 <- sapply(selection3[2:14], mean)  
chemical <- c(2,3,4,5,6,7,8,9,10,11,12,13,14)  
plot(chemical, mean1, col = "red")  
points(chemical, mean2, col="blue", pch="*")  
legend(2,1000,legend=c("Medias en Fabrica 1","Medias en Fabrica 2"),  
      pch=c("o","*"))
```

Podemos realizar scatterplots de la siguiente manera

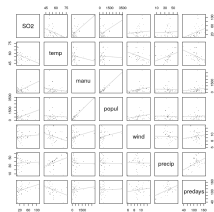
```
plot(popul ~ manu, data = USairpollution)
```

O, dibujar todas las parejas

posibles a la vez, y añadir las rectas de regresión

```
pairs(USairpollution,  
      panel = function(x, y, ...) {  
        points(x, y)      })
```





Implementación Teórica de una DNM

```
DNM <- setRefClass("DNM",  
  fields = list(p = "numeric",  
               media = "matrix",  
               cov = "matrix"))
```

1 Introducción: R

2 R en el análisis multivariante

- Distribución Normal Multivariante. Ejemplos

3 Para ampliar

Computing Machinery and Intelligence Alan Turing (1950)

Artificial Intelligence: A Modern Approach Stuart J. Russell y Peter Norvig

Concrete Problems in AI Safety Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, Dan Mané

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation Miles Brundage, Shahar Avin et al.