

# R Doctorado de Sistemas Complejos

George G. Vega

Superintendencia de Pensiones

April 16, 2013

# Agenda

- 1 Dónde buscar información
- 2 Algunos Paquetes importantes
- 3 Must Know de R
- 4 Algunos ejemplos

# Agenda

- 1 Dónde buscar información
- 2 Algunos Paquetes importantes
- 3 Must Know de R
- 4 Algunos ejemplos

# Dónde buscar información

## Tutoriales Online

- R wiki (oficial) <http://rwiki.sciviews.org/doku.php>
- R-programming (wiki-book)  
[http://en.wikibooks.org/wiki/R\\_Programming](http://en.wikibooks.org/wiki/R_Programming)
- Quick-R <http://www.statmethods.net/>
- Flowing Data <http://flowingdata.com/category/tutorials/>

# Dónde buscar información

## Libros

- Norman Matloff, The Art of R Programming (2009)  
<http://heather.cs.ucdavis.edu/~matloff/132/NSPpart.pdf>
- Introducción a R <http://cran.r-project.org/doc/contrib/R-intro-1.1.0-espanol.1.1.pdf>
- R para principiantes  
[http://cran.r-project.org/doc/contrib/rdebuts\\_es.pdf](http://cran.r-project.org/doc/contrib/rdebuts_es.pdf)

# Dónde buscar información

## Otros

- R-bloggers: Agregador de blogs sobre R  
<http://www.r-bloggers.com/>
- Bioconductor: Herramientas para el análisis genético (biológico)  
<http://www.bioconductor.org/>
- CRAN: The Comprehensive R Archive  
<http://cran.r-project.org>
- R Graph Gallery <http://gallery.r-enthusiasts.com/>
- R Seek: Google de R <http://rseek.org/>

# Agenda

- 1 Dónde buscar información
- 2 Algunos Paquetes importantes
- 3 Must Know de R
- 4 Algunos ejemplos

# Paquetes importantes

**ggplot2** An implementation of the Grammar Graphics (Wickham 2013)

**lattice** Lattice Graphics

**rgl** 3D visualization device system (OpenGL).

**knitr** A general-purpose package for dynamic report generation in R.

**foreign** Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, dBase,

...

**igraph** Network analysis and visualization

**deSolve** General solvers for initial value problems of ordinary differential equations (ODE), partial differential equations (PDE), differential algebraic equations (DAE), and delay differential equations (DDE).



# Agenda

- 1 Dónde buscar información
- 2 Algunos Paquetes importantes
- 3 **Must Know de R**
- 4 Algunos ejemplos

# Must Know de R

- R es un lenguaje de código interpretado (no compilado como C C++ Python, etc). Sigue esquema OOP (Object Oriented Programming), es dinámico (mayor parte de sus chequeos se realiza en ejecución).
- Mayor parte de las funciones de R son transmorficas
- El comportamiento de las funciones dependerá de la clase del objeto (puede ser S3 o S4), a estos se les llama métodos
- para instalar paquetes `install.packages()`
- para pedir ayuda `??'prcomp'`
- Tipos de objetos en R `data.frame`, `list`, `matrix`, `factor`
- Recordar semillas en Procesos Pseudo-Aleatorios (`set.seed()`)
- `lapply` es preferido a `for()` (velocidad).
- R posee varias rutinas matemáticas y estadísticas optimizadas (escritas en C). En el caso de querer más velocidad, se puede integrar con C (`?C`), C++ (`?Call`) Python (`?RPy`).

# Agenda

- 1 Dónde buscar información
- 2 Algunos Paquetes importantes
- 3 Must Know de R
- 4 Algunos ejemplos

# Algunos Ejemplos

## animaciones

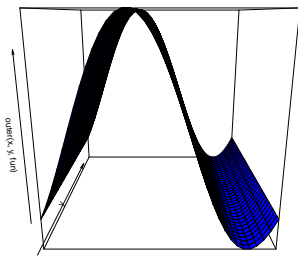
```
library(animation)  
demo("Mandelbrot", echo = FALSE, package = "animation")
```

# Algunos Ejemplos

## gráficos 3D

```
# Funcion
fun <- function(x, y) {
  return(sin(x) - cos(x))
}

# Graficando
x <- y <- seq(0, 2 * pi, pi/18)
persp(x, y, z = outer(x, y, fun), col = "blue")
```



# Algunos Ejemplos

## Análisis de Componentes Principales

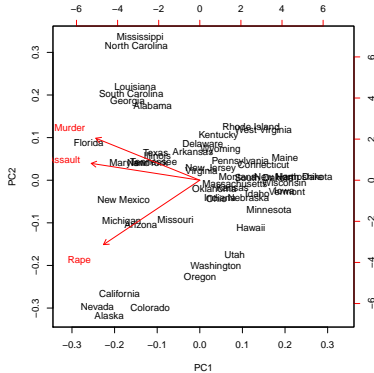
```
data(USArrests)
x <- prcomp(~Murder + Assault + Rape, data = USArrests, scale = TRUE)
summary(x)

## Importance of components:
##              PC1    PC2    PC3
## Standard deviation    1.536 0.677 0.4282
## Proportion of Variance 0.786 0.153 0.0611
## Cumulative Proportion 0.786 0.939 1.0000
```

# Algunos Ejemplos

## Análisis de Componentes Principales

`biplot(x)`



# Algunos Ejemplos

## Modelos lineales (MCO)

```
data(USArrests)
x <- lm(Murder ~ Assault + Rape, data = USArrests)
summary(x)

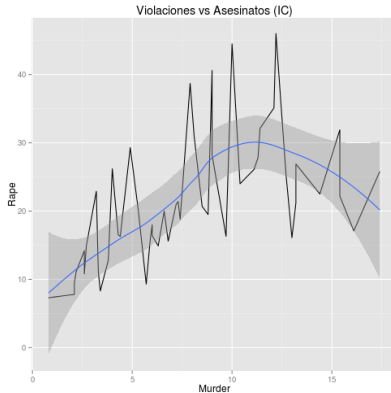
##
## Call:
## lm(formula = Murder ~ Assault + Rape, data = USArrests)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.867 -1.765 -0.375  1.303  7.886
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.41886    0.97618   0.43    0.67
## Assault      0.04003    0.00609   6.58 3.6e-08 ***
## Rape         0.02514    0.05416   0.46    0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.65 on 47 degrees of freedom
## Multiple R-squared:  0.645, Adjusted R-squared:  0.63
## F-statistic: 42.6 on 2 and 47 DF,  p-value: 2.76e-11
```



# Algunos Ejemplos

## ggplot2

```
library(ggplot2)
m <- ggplot(USArrests, aes(x=Murder,y=Rape)) # Def del objeto
m + geom_line() + # Grafico de linea
  geom_smooth(method="loess") + # IC
  labs(title="Violaciones vs Asesinatos (IC)") # Etiqueta
```



# R Doctorado de Sistemas Complejos

George G. Vega

Superintendencia de Pensiones

April 16, 2013

(presentación creada en R + knitr +  $\LaTeX$ )