

PRÁCTICAS DE ORDENADOR II

En los ejercicios que siguen, se puede usar **EXCEL** o **R**.

[1] Realizar un análisis de componentes principales para las dos primeras columnas del conjunto de datos `ventas.txt`, disponible en `Datos.txt.zip`. Distinguir entre los casos con o sin tipificación, indicando cuál es el **más adecuado** en esta situación.

[2] Realizar un análisis de componentes principales para el conjunto de datos `sctkpr.txt`, que se encuentra disponible en `Datos.txt.zip`. Para ello, multiplicar los datos por 100 o, equivalentemente, **expresar los retornos semanales en forma de porcentaje**. Si el nuevo conjunto de datos se llama, por ejemplo, `datos1.*`:

- a) Calcular las desviaciones típicas de las columnas.
- b) **Analizar y comparar** los resultados obtenidos a partir de las matrices de covarianzas y correlaciones muestrales de `datos1.*`.

[3] Considerar los retornos asociados a los **valores diarios de cierre** del índice bursátil Dow Jones durante los años 1992-1998 para las empresas

```
I = c('MO', 'KO', 'EK', 'HWP', 'INTC', 'MSFT', 'IBM', 'MCD', 'WMT', 'DIS') ,
```

donde 'MO': Philip Morris; 'KO': Coca Cola; 'EK': Kodak; 'HWP': Hewlett Packard; 'INTC': Intel; 'MSFT': Microsoft; 'MCD': McDonald's; 'WMT': Walmart; y 'DIS': Disney. Los datos están disponibles como parte del conjunto `dowjones.txt` en el enlace `Datos.txt.zip`. No obstante, estos se pueden obtener directamente con **R** usando el siguiente código, cuya ejecución requiere de la **instalación preliminar del paquete QRM**:

```
> library(QRM); data(DJ);
> I = c('MO', 'KO', 'EK', 'HWP', 'INTC', 'MSFT', 'IBM', 'MCD', 'WMT', 'DIS');
> data.0 = window(DJ[,I], '1992-01-01', '1998-12-31');
> data.1 = 100*data.frame(returns(data.0, method = 'discrete'));
> write.table(round(data.1, d = 4), 'dowjones10.txt');
```

El fichero `dowjones10.txt`, que se encontrará localizado en el directorio de trabajo, contiene los datos de interés. Se trata entonces de realizar un análisis de componentes principales usando la matriz de covarianzas de `data.1`. El objetivo es comprobar que las **dos fuentes de variación dominantes** son un **factor global de mercado**, y una **comparación** de los valores **tecnológicos**, `c('HWP', 'INTC', 'MSFT', 'IBM')`, y de **consumo**, `c('MO', 'KO', 'EK', 'MCD', 'WMT', 'DIS')`.

Comentario. Para realizar esta práctica es conveniente familiarizarse primero, además de con el contenido de `apexcel.txt`, con los programas de **R**

`opclR.txt` `plot_opclR.txt` ,

del enlace `Software2.zip`. Para ello, se puede seguir el siguiente procedimiento:

- a) Descargar el fichero `openc1.txt`, disponible en el enlace `Datos.txt.zip`.

- b) Cambiar el directorio de la sesión de **Rstudio** al lugar en el que **esté almacenado** `openc1.txt`. Usar para ello la instrucción

`'Set Working Directory' ,`

o bien directamente `'Control + Mayusc + H'`.

- c) Activar y limpiar la consola, `'Control + L'`, y pegar el contenido de los ficheros `opclR.txt` (`plot_opclR.txt`).
- d) Pulsar **Enter**, y analizar el procedimiento empleado y los resultados obtenidos.

Para sucesivas ejecuciones, es de utilidad efectuar `graphics.off()`, con el fin de limpiar los gráficos, y `rm(list = ls())`, para eliminar los objetos creados anteriormente.