

# Fundamentos de Análisis de Datos

## **Análisis de Componentes Principales**

Dra. Andrea Alejandra Rey

Especialización en Ciencia de Datos - ITBA



Conservación de información relevante

Reducción de la dimensión

Conjunto de datos con muchas variables

# Análisis de componentes principales

- ➡ Es conocido por el acrónimo PCA, debido a su equivalente en inglés *Principal Component Analysis*.
- ➡ Es una técnica de análisis de datos multivariados basado en métodos de proyección.
- ➡ Es un método utilizado para reducir la dimensionalidad de grandes conjuntos de datos, transformando un gran conjunto de variables en uno más pequeño que contiene la mayor parte de la información de los datos.
- ➡ Es evidente que los conjuntos de datos más pequeños son más fáciles de explorar, visualizar y analizar. Es por ello que la reducción del número de variables resulta una idea muy atractiva. Sin embargo, esta reducción se produce a expensas de la precisión.

## Pregunta

¿Cómo podemos reducir la dimensionalidad, por simplicidad, sin perder demasiada precisión?

# Pasos del PCA

- 1 Estandarización de las variables.
- 2 Cálculo de la matriz de covarianzas.
- 3 Cálculo de autovalores y autovectores.
- 4 Creación de un vector de características con las componentes principales.
- 5 Proyección de los datos en los ejes de las componentes principales.

# ESTANDARIZACIÓN

# Estandarización

- ▶ PCA es sensible a las varianzas de las variables iniciales.
- ▶ Aquellas variables cuyos valores pertenezcan a rangos más amplios, serán más dominantes y producirán resultados sesgados.
- ▶ Las variables deben estandarizarse con el fin de que cada una de ella contribuya de manera equitativa en al análisis.
- ▶ Luego de la estandarización, todas las variables se transforman a la misma escala.

Recordemos que si  $x = \{x_1, x_2, \dots, x_n\}$  es una muestra de una variable aleatoria  $X$ , la muestra estandarizada está formada por los elementos:

$$\frac{x_i - \bar{x}}{s_x},$$

para  $i = 1, 2, \dots, n$ .

# MATRIZ DE COVARIANZAS

# Matriz de covarianzas

- ➡ Algunas veces las variables están altamente correlacionadas de manera tal que contienen información redundante.
- ➡ Para poder identificar las relaciones entre las variables, se calcula la matriz de covarianzas.

Recordemos que la matriz de covarianzas de  $X_1, X_2, \dots, X_k$  es una matriz simétrica y semidefinida positiva, de la forma:

$$\Sigma = \begin{pmatrix} \text{Var}(X_1) & \text{Covar}(X_1, X_2) & \cdots & \text{Covar}(X_1, X_k) \\ \text{Covar}(X_2, X_1) & \text{Var}(X_2) & \cdots & \text{Covar}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Covar}(X_k, X_1) & \text{Covar}(X_k, X_2) & \cdots & \text{Var}(X_k) \end{pmatrix}$$



# AUTOVALORES Y AUTOVECTORES



Recurriendo al Álgebra Lineal...

# Determinante de una matriz

El **determinante** es un número real que sólo puede calcularse para matrices cuadradas, donde la cantidad de filas coincide con la cantidad de columnas.

$$A \in \mathbb{R}^{2 \times 2}$$

$$\det(A) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$A \in \mathbb{R}^{3 \times 3}$$

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

# Determinante de una matriz

## Observación

- ➡ Para calcular el determinante de una matriz de  $3 \times 3$  es necesario calcular 3 determinantes de  $2 \times 2$ .
- ➡ De manera similar, se puede calcular el determinante de una matriz de  $4 \times 4$  mediante 4 determinantes de  $3 \times 3$ .
- ➡ También, se puede calcular el determinante de una matriz de  $5 \times 5$  usando 5 determinantes de  $4 \times 4$ .
- ➡ Y de manera inductiva, definir el determinante de una matriz de  $n \times n$ .

# Determinante de una matriz

## Relación del determinante con los sistemas lineales homogéneos

Sea el sistema lineal homogéneo  $AX = 0$  con  $n$  incógnitas y  $n$  ecuaciones; es decir:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Observar que este sistema siempre tiene solución.

Además,

El sistema admite a  $(0, 0, \dots, 0)$  como única solución  $\iff \det(A) \neq 0$

El sistema admite infinitas soluciones  $\iff \det(A) = 0$

# Autovalores y autovectores

## Definición

Sea  $A \in \mathbb{R}^{n \times n}$ . Se dice que  $\lambda \in \mathbb{R}$  es un **autovalor** de  $A$  si existe un  $v = (v_1, v_2, \dots, v_n)' \in \mathbb{R}^n$  no nulo tal que:

$$Av = \lambda v.$$

En este caso, se dice que  $v$  es un **autovector** de  $A$  de autovalor  $\lambda$ .

- ➡ También se utiliza la terminología **valor propio** y **vector propio**.
- ➡ Sus equivalentes en inglés son *eigenvalue* y *eigenvector*.

# Autovalores y autovectores

Observemos lo siguiente:

$$Av = \lambda v$$

$$Av - \lambda v = 0$$

$$(A - \lambda I)v = 0 \longrightarrow \text{Es un sistema lineal homogéneo}$$

En la definición de autovalor se pide que  $v$  no sea el vector nulo, lo que implica que el sistema tiene infinitas soluciones y por ende:

$$\boxed{\det(A - \lambda I) = 0}.$$

# Autovalores y autovectores

## Polinomio característico

Sea  $A \in \mathbb{R}^{n \times n}$ . Su **polinomio característico** es un polinomio de grado  $n$  definido como:

$$\chi_A(\lambda) = \det(A - \lambda I).$$

## Observación

Los autovalores de una matriz son las raíces de su polinomio característico.



# Autovalores y autovectores

Consideremos como ejemplo a la matriz  $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ .

Para hallar sus autovalores debemos calcular su polinomio característico:

$$\begin{aligned}\chi_A(\lambda) &= \det(A - \lambda I) = \det \left[ \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right] \\ &= \begin{vmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 - 4.\end{aligned}$$

Ahora hallamos sus raíces:

$$\begin{aligned}(1 - \lambda)^2 - 4 &= 0 \iff (1 - \lambda)^2 = 4 \\ &\iff 1 - \lambda = 2 \text{ o } 1 - \lambda = -2 \\ &\iff \boxed{\lambda = -1} \text{ o } \boxed{\lambda = 3}.\end{aligned}$$

Luego, los autovalores de  $A$  son -1 y 3.

# Autovalores y autovectores

Para hallar los autovectores de  $A$ , debemos analizar cada autovalor por separado.

⇒ Para  $\lambda = -1$ , debemos resolver el sistema:

$$(A+I)v = 0 \iff \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} 2v_1 + 2v_2 = 0 \\ 2v_1 + 2v_2 = 0 \end{cases} \implies v_2 = -v_1.$$

Tenemos que  $\boxed{(1, -1)'}$  es un autovector de  $A$  de autovalor  $-1$ .

⇒ Para  $\lambda = 3$ , debemos resolver el sistema:

$$(A-3I)v = 0 \iff \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{cases} -2v_1 + 2v_2 = 0 \\ 2v_1 - 2v_2 = 0 \end{cases} \implies v_2 = v_1.$$

Tenemos que  $\boxed{(1, 1)'}$  es un autovector de  $A$  de autovalor  $3$ .

# Autovalores y autovectores

Observemos lo siguiente:

$$\|(1, -1)\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} \implies \left( \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) \text{ es autovector de } A \text{ de}$$

$\lambda = -1$  con norma 1

$$\|(1, 1)\| = \sqrt{1^2 + 1^2} = \sqrt{2} \implies \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \text{ es autovector de } A \text{ de}$$

$\lambda = 3$  con norma 1

Estos nuevos autovectores de norma igual a 1 se dicen **normalizados**.

# Autovalores y autovectores

Consideremos finalmente las siguientes matrices:

⇒  $D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$ , la matriz diagonal que tiene en su diagonal principal a los autovalores de  $A$

⇒  $Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$ , la matriz que tiene por columnas a los autovectores de  $A$  normalizados.

Observemos que:

$$QQ' = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Hemos probado que  $Q' = Q^{-1}$ , lo cual dice que la matriz es **ortogonal**. Esto implica que sus columnas son vectores ortogonales (perpendiculares).

# Autovalores y autovectores

Entonces:

$$\begin{aligned} QDQ' &= \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}' \\ &= \begin{pmatrix} \frac{-\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{3\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 2 & 1 \end{pmatrix} = A. \end{aligned}$$

Pregunta

¿Casualidad?

# Autovalores y autovectores

## Teorema

Sea  $A \in \mathbb{R}^{n \times n}$  una matriz simétrica. Entonces  $A$  puede descomponerse como

$$A = QDQ',$$

donde  $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$  es la matriz diagonal formada con los autovalores de  $A$  y  $Q = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$  es una matriz ortogonal cuyas columnas corresponden a autovectores de  $A$  normalizados.

## Observación

Bajo las condiciones del teorema:

$$A = QDQ' \iff Q'AQ = D.$$

# Autovalores y autovectores

## Teorema

Si  $A \in \mathbb{R}^{n \times n}$  una matriz semidefinida positiva, entonces:

$$\max_{x: \|x\|=1} x'Ax = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = \min_{x: \|x\|=1} x'Ax,$$

donde  $\lambda_1, \lambda_2, \dots, \lambda_n$  son los autovalores de  $A$ .

# Autovalores y autovectores

## Observación

- ➡ Un autovector siempre está asociado a un autovalor, por lo que estos conceptos siempre vienen de a pares.
- ➡ Existen tantos autovalores como la dimensión del conjunto de variables de los datos.
- ➡ Veremos que los autovectores de la matriz de covarianzas son las direcciones de las componentes principales.
- ➡ Veremos que los autovalores de la matriz de covarianzas representan la cantidad de varianza que lleva cada componente principal.



## Ejemplo\*



---

\*Base de datos disponible en <https://archive.ics.uci.edu/ml/datasets/Student+Performance>

# COMPONENTES PRINCIPALES

# Componentes principales

- ➡ Las nuevas variables se construyen como combinación lineal de las variables iniciales.
- ➡ Estas combinaciones se realizan de manera tal de evitar la correlación, buscando direcciones ortogonales (perpendiculares).
- ➡ Existen tantas componentes principales como variables iniciales.
- ➡ La mayoría de la información de las variables iniciales estará concentrada en la primera componente, mientras que la mayoría del resto de la información se encontrará en la segunda componente, y así sucesivamente.
- ➡ Esta manera de organizar la información permite descartar aquellas componentes con muy poca información y conservar el resto de las componentes como las nuevas variables, reduciendo de este modo la dimensión.
- ➡ Las componentes principales son menos interpretables puesto que no tienen un significado real.

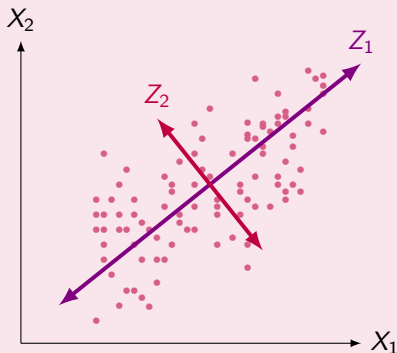
# Componentes principales

## Interpretación

- ➡ Las componentes principales representan las direcciones de los datos que explican una cantidad de varianza máxima; es decir, las rectas que captan la mayor información.
- ➡ Cuanto mayor sea la varianza de una recta, mayor será la dispersión de los puntos a lo largo de ella, y por ende tendrá más información.
- ➡ Las componentes principales son los nuevos ejes que permiten un mejor ángulo de visualización para notar diferencias entre las observaciones.

# Componentes principales

Supongamos que tenemos dos variables  $X_1$  y  $X_2$  y datos distribuidos como se muestra a continuación.



- ➡  $Z_1$  sigue la dirección en la cual las observaciones varían más.
- ➡  $Z_2$  sigue la segunda dirección en la que los datos muestran mayor varianza y que resulta perpendicular a la primera dirección.

# Componentes principales

Consideramos un conjunto de características  $X_1, X_2, \dots, X_k$ .

Vamos a estudiar un promedio ponderado, dado por la **combinación lineal**:

$$\phi_1 X_1 + \phi_2 X_2 + \dots + \phi_k X_k,$$

donde  $\phi_1, \phi_2, \dots, \phi_k \in \mathbb{R}$ .

Necesitamos además que esta combinación lineal esté **normalizada**, es decir:

$$\|\phi\|^2 = \sum_{i=1}^k \phi_i^2 = 1.$$

El vector  $\phi = (\phi_1, \phi_2, \dots, \phi_k)$  se llama **vector de pesos** o **vector de cargas** (*loadings*, en inglés).

# Componentes principales

## Pregunta

¿Cómo elegimos el vector  $\phi$ ?

Indicamos la combinación lineal normalizada como  $\phi' \mathbf{X}$ .

El objetivo es maximizar su varianza; es decir, hallar:

$$\phi = \arg \max_{\{\psi: \|\psi\|=1\}} \text{Var}(\psi' \mathbf{X}) = \arg \max_{\{\psi: \|\psi\|=1\}} \psi' \text{Var}(\mathbf{X}) \psi,$$

donde  $\text{Var}(\psi' \mathbf{X})$  denota la matriz de covarianzas de  $\psi' \mathbf{X}$ .

Como la matriz de covarianzas es semidefinida positiva, podemos aplicar un teorema previo para deducir que  $\phi$  es el autovector de la matriz de covarianzas asociado al mayor de sus autovalores.

# Componentes principales

## Pregunta

¿Cómo calculamos las componentes principales?

Sea  $\mathbf{X}$  un vector aleatorio de  $k$  componentes, con media  $E(\mathbf{X}) = \boldsymbol{\mu}$  y matriz de covarianzas  $\boldsymbol{\Sigma} = QDQ'$ .

Las componentes principales están dadas por:

$$\mathbf{Y} = Q'(\mathbf{X} - \boldsymbol{\mu}).$$

## Media de las componentes principales

$$E(\mathbf{Y}) = E[Q'(\mathbf{X} - \boldsymbol{\mu})] = Q'E(\mathbf{X} - \boldsymbol{\mu}) = Q' \underbrace{E(\mathbf{X})}_{\boldsymbol{\mu}} - \underbrace{E(\boldsymbol{\mu})}_{\boldsymbol{\mu}} \implies E(\mathbf{Y}) = \mathbf{0}$$



# Componentes principales

## Varianzas y covarianzas de las componentes principales

Para  $i, j = 1, 2, \dots, k$ :

$$\begin{aligned}\text{Covar}(Y_i, Y_j) &= \text{Covar}[Q'_i(\mathbf{X} - \boldsymbol{\mu}), Q'_j(\mathbf{X} - \boldsymbol{\mu})] = Q'_i \text{Covar}(\mathbf{X} - \boldsymbol{\mu}) Q_j \\ &= Q'_i \text{Covar}(\mathbf{X}) Q_j = Q'_i \Sigma Q_j = D_{ij} = \begin{cases} \lambda_i & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}\end{aligned}$$

Entonces:

$$\boxed{\text{Var}(Y_i) = \lambda_i}$$

y para  $i \neq j$ :

$$\boxed{\text{Covar}(Y_i, Y_j) = 0}.$$

# Componentes principales

Observar que:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \implies \text{Var}(Y_1) \geq \text{Var}(Y_2) \geq \cdots \geq \text{Var}(Y_k).$$

$Y_1$  es la primera componente principal

$Y_2$  es la segunda componente principal

$\vdots$

$Y_k$  es la  $k$ -ésima componente principal

# PROYECCIÓN

# Proyección

- ➡ Los pasos explicados hasta el momento no producen ningún cambio en los datos, que siguen estando expresados en términos de las variables iniciales.
- ➡ El vector de características reorienta los datos desde los ejes originales hacia los ejes representados por las componentes principales.

# Proyección

Si tenemos una muestra de un vector aleatorio  $\mathbf{X} = (X_1, X_2, \dots, X_k)$ , el primer paso es estandarizar los datos.

Denotamos por  $\mathbf{x}$  a la muestra estandarizada.

Podemos calcular  $\hat{\Sigma}$  la matriz de covarianzas muestral, con su descomposición:

$$\hat{\Sigma} = \mathbf{Q}'\mathcal{D}\mathbf{Q},$$

donde  $\mathbf{Q}$  es una matriz ortogonal cuyas columnas  $v_1, v_2, \dots, v_k$  son autovectores de  $\hat{\Sigma}$  correspondientes a los autovalores  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ , y  $\mathcal{D}$  es la matriz diagonal con estos autovalores en su diagonal.

# Proyección

Entonces, para  $j = 1, 2, \dots, k$ , las componentes principales están dadas por:

$$Y_j = (\mathbf{X} - \bar{\mathbf{x}})v_j,$$

donde  $\bar{\mathbf{x}} = 0$  por la estandarización.

Las proyecciones de los datos sobre las componentes principales se calculan como:

$$\text{proy}_{Y_j}(\mathbf{x}_i) = \mathbf{x}_i v_j.$$

# Explicación de la varianza

## Pregunta

¿Cuanta información es capaz de capturar cada una de las componentes principales obtenidas?

## Proporción de varianza explicada

$$\text{PVE}_q = \frac{\sum_{i=1}^q \text{Var}(Y_i)}{\sum_{i=1}^k \text{Var}(Y_i)} = \frac{\sum_{i=1}^q \lambda_i}{\sum_{i=1}^k \lambda_i}$$

## Observación

Las primeras  $q$  componentes principales explican el  $\text{PVE}_q 100\%$  de la variación.

# Cantidad de componentes principales

## Vector de características

Es una matriz que tiene por columnas aquellos autovectores de las componentes principales elegidas.

## Pregunta

¿Cómo elegimos la cantidad de componente principales con las cuales trabajar?

Equivalentemente, ¿cuántas componentes vamos a descartar?



# Métodos para elegir la cantidad de características

## Método I

Selección aleatoria. Por ejemplo, si se desea una mejor visualización en dos dimensiones, se eligen las dos primeras características.

## Método II

Calcular la proporción de varianza explicada por cada característica, elegir un umbral y agregar características hasta alcanzar ese umbral.

## Método III

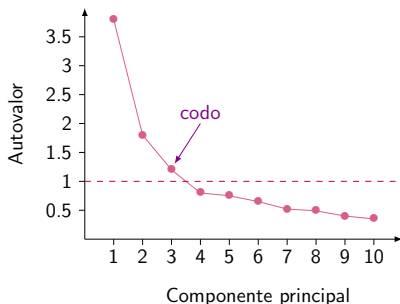
Usar un *scree plot* para identificar el punto para el cual agregar una nueva característica implica una caída significativa en la varianza explicada en relación con la característica anterior y elegir la cantidad de características a incluir hasta ese punto.

# Scree plot

- ➡ En el eje vertical se muestran los autovalores correspondientes a cada una de las componentes principales representadas en el eje horizontal.
- ➡ Todos presentan una curva similar, que comienza con un valor alto en la izquierda, decrece rápidamente hasta un punto donde se empieza a estabilizar.
- ➡ El punto donde la curva empieza a aplanarse se llama “codo”.

# Scree plot

## Primer ejemplo

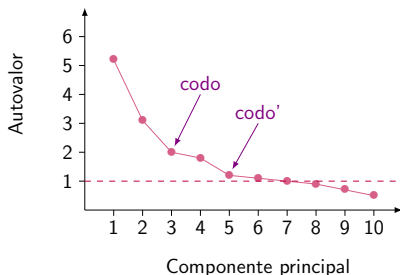


La regla de Kaiser establece conservar las componentes principales con autovalores al menos 1.

En el ejemplo, nos quedaríamos con las primeras tres componentes principales.

# Scree plot

## Segundo ejemplo



Según la regla de Kaiser nos quedaríamos con las primeras siete componentes principales.

Pero, ¿es ésta la mejor opción?

El scree plot sugiere quedarnos con las primeras tres, o a lo sumo cinco, componentes principales.

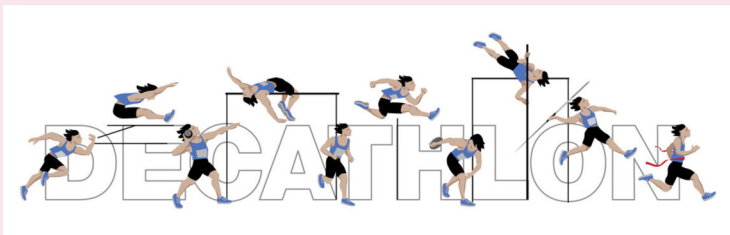
# Interpretación del PCA

Es un método que proporciona la siguiente información:

- ▢ **Matriz de covarianzas:** una medida de cómo cada variable se relaciona con cada una de las demás variables.
- ▢ **Autovectores:** direcciones en las cuales se dispersan los datos.
- ▢ **Autovalores:** importancia de estas direcciones.

PCA combina las variables predictoras permitiendo descartar aquellos autovectores que proporcionan una cantidad de información despreciable.

## Ejemplo\*



\*Base de datos decathlon2 del paquete factoextra de R

# Regresión por componentes principales

- ➡ A menudo, un pequeño número de componentes principales es suficiente para explicar la mayor parte de la variabilidad de los datos, así como la relación con la variable respuesta.
- ➡ Esta suposición no está garantizada, pero a menudo es lo suficientemente razonable como para dar buenos resultados.

## PCR

La **regresión por componentes principales**, conocida como PCR (*Principal Component Regression*) consiste en hallar las  $q$  primeras componentes principales y usarlas como predictores en un modelo de regresión lineal ajustado por mínimos cuadrados.

# Regresión por componentes principales

## Ventajas 😊

- ➡ Puede mejorar el rendimiento del modelo predictivo al resolver problemas causados por la multicolinealidad o el gran número de variables.
- ➡ Es un procedimiento simple de dos pasos: reducción de la dimensión y ajuste por mínimos cuadrados.

## Desventajas 😞

- ➡ No realiza la selección de características.
- ➡ Carece de interpretación puesto que las estimaciones se realizan sobre las componentes principales que no tienen un significado en el modelo.
- ➡ Se basa en la suposición de que las direcciones en las cuales los predictores muestran la mayor variación son las que predicen la respuesta, lo cual no siempre es cierto.



# Ejemplo\*



---

\*Base de datos crimeData del paquete mogavs de R