Idea del Análisis Discriminante Lineal

LDA en dos palabras

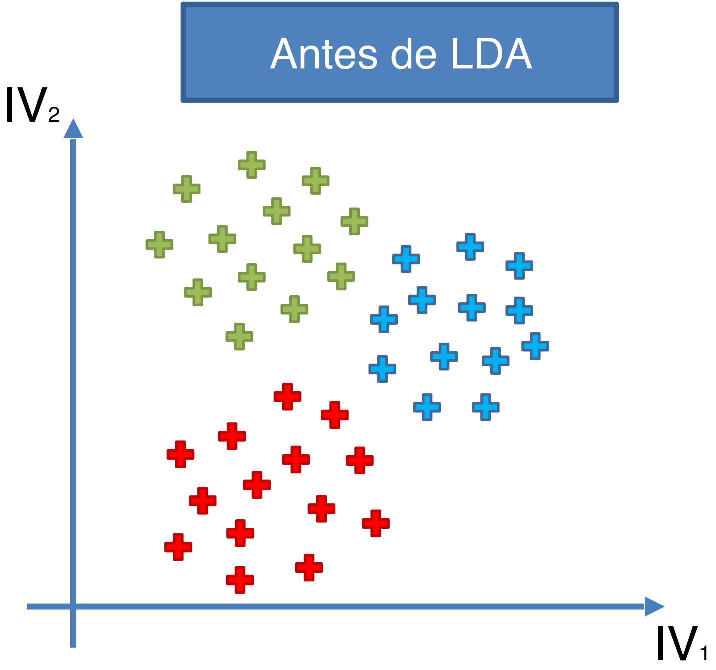
De las n variables independientes del dataset, LDA extrae las $p \leq n$ nuevas variables independientes que separan la mayoría de clases de la variable dependiente.

LDA en dos palabras

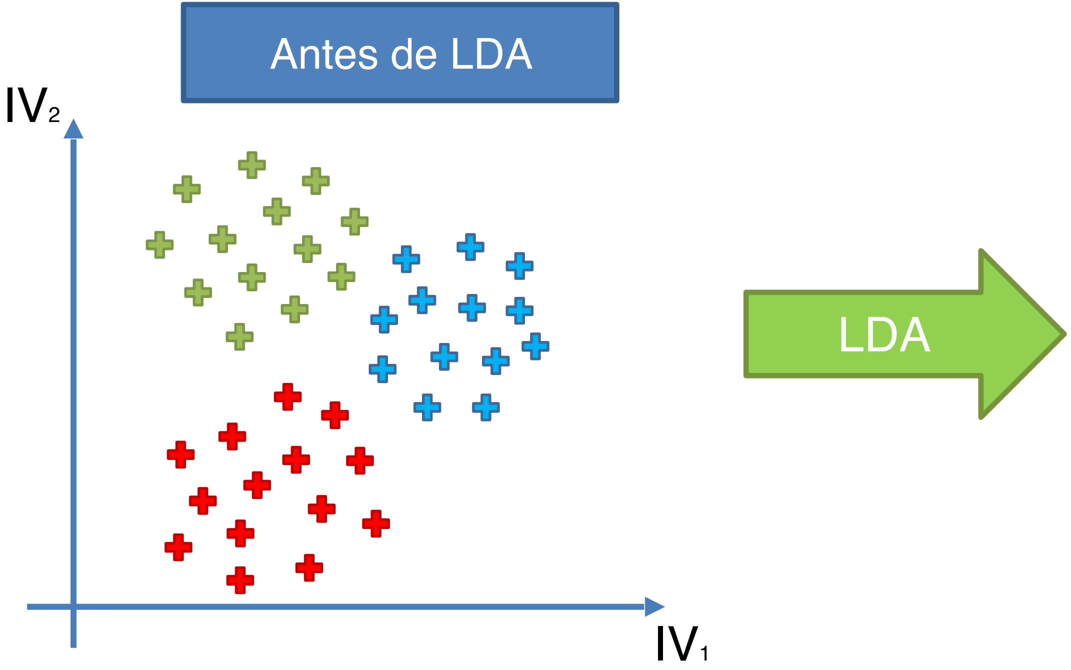
De las n variables independientes del dataset, LDA extrae las $p \leq n$ nuevas variables independientes que separan la mayoría de clases de la variable dependiente.

→ Como se usa la VD en el modelo, el LDA resulta ser un modelo supervisado.

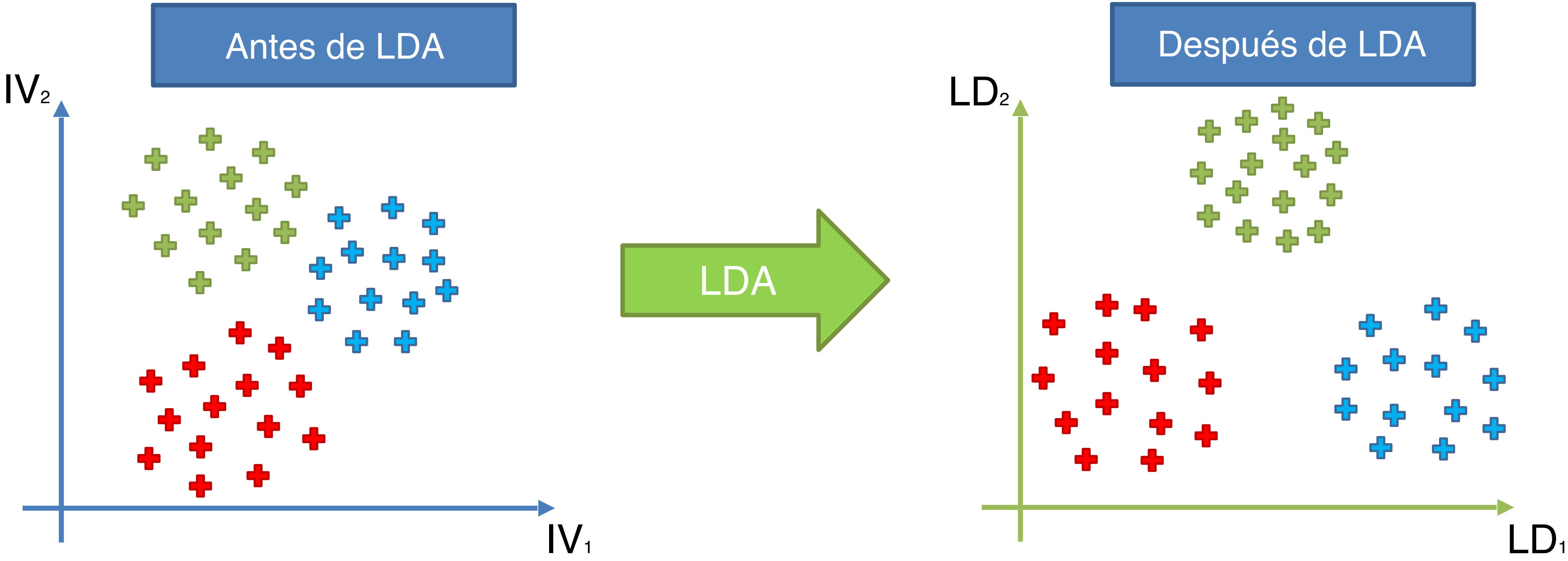
LDA encuentra las direcciones de máxima separación de clases



LDA encuentra las direcciones de máxima separación de clases



LDA encuentra las direcciones de máxima separación de clases



LD_1 y LD_2 son las direcciones de máxima separación de clases

Las matemáticas tras el LDA

PASO 1: Aplicar escalado de variables a la matriz de características X, compuesta por n variables independientes.

Las matemáticas tras el LDA

PASO 2: Sea C el número de clases, calcular C vectores m-dimensionales, de modo que cada uno contenga las medias de las características de las observaciones para cada clase.

Ex: supongamos que las VD tienen dos clases 0 y 1, y sea x_j^i la característica j-ésima de la observación i-ésima, entonces

$$\boldsymbol{\mu_0} = \left| \begin{array}{c} \frac{1}{n_0} \sum\limits_{\substack{i=1,\dots,n \\ y^i \in \text{class } 0}} x_1^i \\ \vdots \\ \frac{1}{n_0} \sum\limits_{\substack{i=1,\dots,n \\ y^i \in \text{class } 0}} x_m^i \end{array} \right. \qquad \boldsymbol{\mu_1} = \left| \begin{array}{c} \frac{1}{n_1} \sum\limits_{\substack{i=1,\dots,n \\ y^i \in \text{class } 1}} x_1^i \\ \vdots \\ \frac{1}{n_1} \sum\limits_{\substack{i=1,\dots,n \\ y^i \in \text{class } 1}} x_m^i \end{array} \right.$$

Las matemáticas tras el LDA

PASO 3: Calculamos la matriz de productos cruzados centrados en la media para cada clase, que mide la varianza dentro de cada clase

Con nuestro ejemplo de las clases 0 y 1, las dos matrices de productos cruzados S_0 y S_1 para las respectivas clases 0 y 1 son:

$$S_0 = \sum_{\substack{i=1,\ldots,n \\ y^i \in \text{class } 0}} \left((x_1^i, \ldots, x_m^i) - \boldsymbol{\mu_0}\right)\left((x_1^i, \ldots, x_m^i) - \boldsymbol{\mu_0}\right)^T$$

$$S_1 = \sum_{\substack{i=1,\ldots,n \\ y^i \in \text{class } 1}} \left((x_1^i, \ldots, x_m^i) - \boldsymbol{\mu_1}\right)\left((x_1^i, \ldots, x_m^i) - \boldsymbol{\mu_1}\right)^T$$

Las matemáticas tras el LDA

PASO 4: Calculamos la covarianza normalizada de todas matrices anteriores, W

Con nuestro ejemplo de las clases 0 y 1, la covarianza normalizada W es simplemente:

$$W = \frac{1}{n_0} S_0 + \frac{1}{n_1} S_1$$

Las matemáticas tras el LDA

PASO 5: Calculamos la matriz de covarianza global entre clases, B

Con nuestro ejemplo de las clases 0 y 1, la matriz de covarianza global entre clases B es simplemente:

$$B = n_0(\boldsymbol{\mu_0} - \boldsymbol{\mu})(\boldsymbol{\mu_0} - \boldsymbol{\mu})^T + n_1(\boldsymbol{\mu_1} - \boldsymbol{\mu})(\boldsymbol{\mu_1} - \boldsymbol{\mu})^T$$

$$\text{where } \boldsymbol{\mu} = (\underbrace{\mu, \ldots, \mu}_{m \text{ times}})^T \text{ with } \mu = \frac{1}{n} \sum_{i=1,\ldots,n} \sum_{j=1,\ldots,m} x_j^i$$

Las matemáticas tras el LDA

PASO 6: Calculamos los valores y vectores propios de la matriz

$$W^{-1}B$$

Las matemáticas tras el LDA

PASO 7: Elegimos los p valores propios más grandes como el número de dimensiones reducidas

Las matemáticas tras el LDA

PASO 8: Los p vectores propios asociados a los p valores propios más grandes son los discriminantes lineales. El espacio m-dimensional del dataset original se proyecta al nuevo subespacio p-dimensional de características, aplicando la matriz de proyecciones (que tiene los p vectores propios por columnas).