PCA → Principal Components Analysis [Dimensionality Reduction]

X = tipos de carros $\left[\text{vel max, radio de giro, rpm, c.f. } \dots\right]_n$

x_i, x_j
- x_i = vel max (km/h)
- x_j = vel max (millas/h)

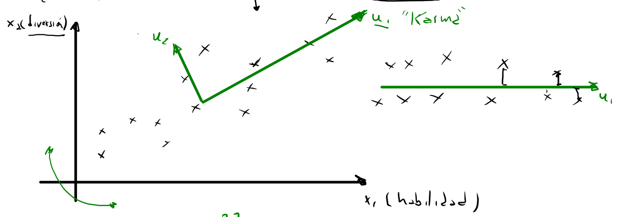
} redundante linealmente dependientes

$\underline{x_i = K \cancel{x_j} + c}$

X = encuesta a pilotos de drones.

$x_1^{(i)}$ = habilidad del piloto
$x_2^{(i)}$ = disfruta del hobbie

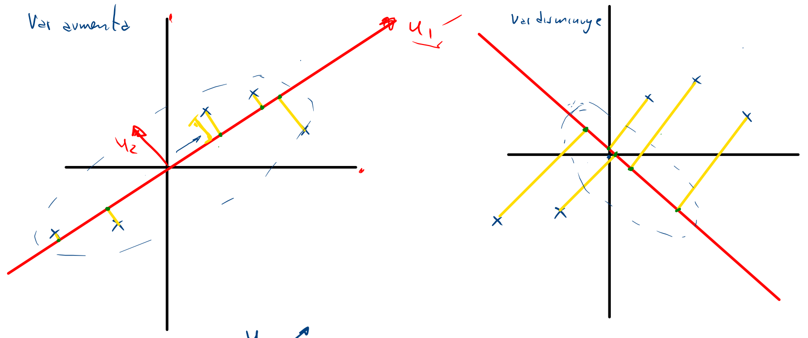
$\underline{x_1 = K x_2 + c}$



PCA → $\underline{u_1 = ??}$

* Normalizar datos (!)

$$x_j^{(i)} \leftarrow \frac{x_j^{(i)} - \mu_j}{\sigma_j}$$



u

x → dist, proy

$x^T u$ → cuan alejado está x del origen

$\|u\|_2 = 1$

$$\underbrace{\frac{1}{n}\sum_{i=1}^{n} \left(x^{(i)T} u\right)^2}_{\text{varianza}} = \frac{1}{n}\sum_{i}^{n} u^T x^{(i)} x^{(i)T} u$$

$$= u^T \left(\frac{1}{n}\sum x^{(i)} x^{(i)T}\right) u$$

$\Sigma = \frac{1}{n}\sum_{i}^{n} x^{(i)} x^{(i)T} \Rightarrow$ matriz de covarianza de X

→ $\underline{\text{autovectores}}$

→ $\underline{\text{PCA}}$

u_j = autovector de Σ

De forma general:

$x^{(i)} \in \mathbb{R}^d$ $\quad \underline{k < d}$ $\quad \Sigma \to u_1 \dots u_d$

$y^{(i)} \in \mathbb{R}^k$ $\quad \to \underline{u_1 \dots u_k}$ | PC

$$y^{(i)} = \begin{bmatrix} u_1^T x^{(i)} \\ u_2^T x^{(i)} \\ \vdots \\ u_k^T x^{(i)} \end{bmatrix} \in \mathbb{R}^k$$

Ejemplos:

- Compresión (con pérdida)
- Visualización de datos *
- Preprocesamiento de datos
- Reducción de ruido *