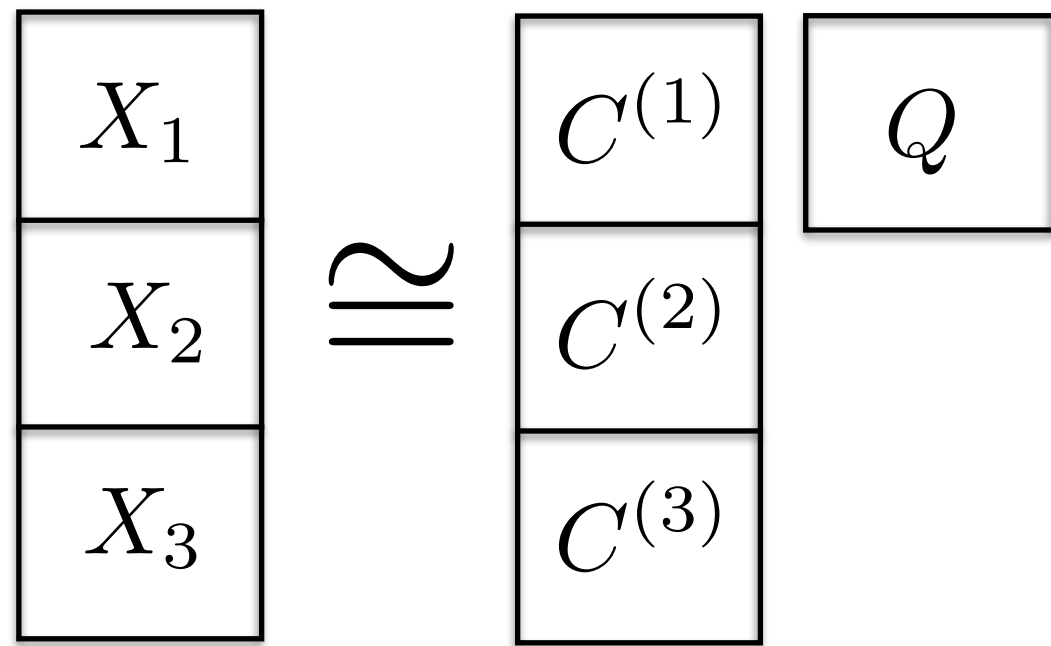


k -means:



Columns of Q are
indicator vectors

A new vector in the i -th language,
 $x \in \mathbb{R}^{n_i}$ is mapped to α - the new
coordinates minimize

$$||X_i - C^{(i)}\alpha||$$

This has the solution

$$\alpha = \underbrace{\left(C^{(i)T}C^{(i)}\right)^{-1}C^{(i)T}}_{P_i}x$$