Bag of Words ➜ Vector Space Model (VSM)

pseudo doc ("**cat**")

$\mathbf{word_1}$

$\mathbf{d1}=(x_1, \ldots x_N)$

N: vocabulary size

$\mathbf{d2}=(y_1, \ldots y_N)$

pseudo doc ("**dog**")

$\mathbf{word_N}$

$\mathbf{word_2}$
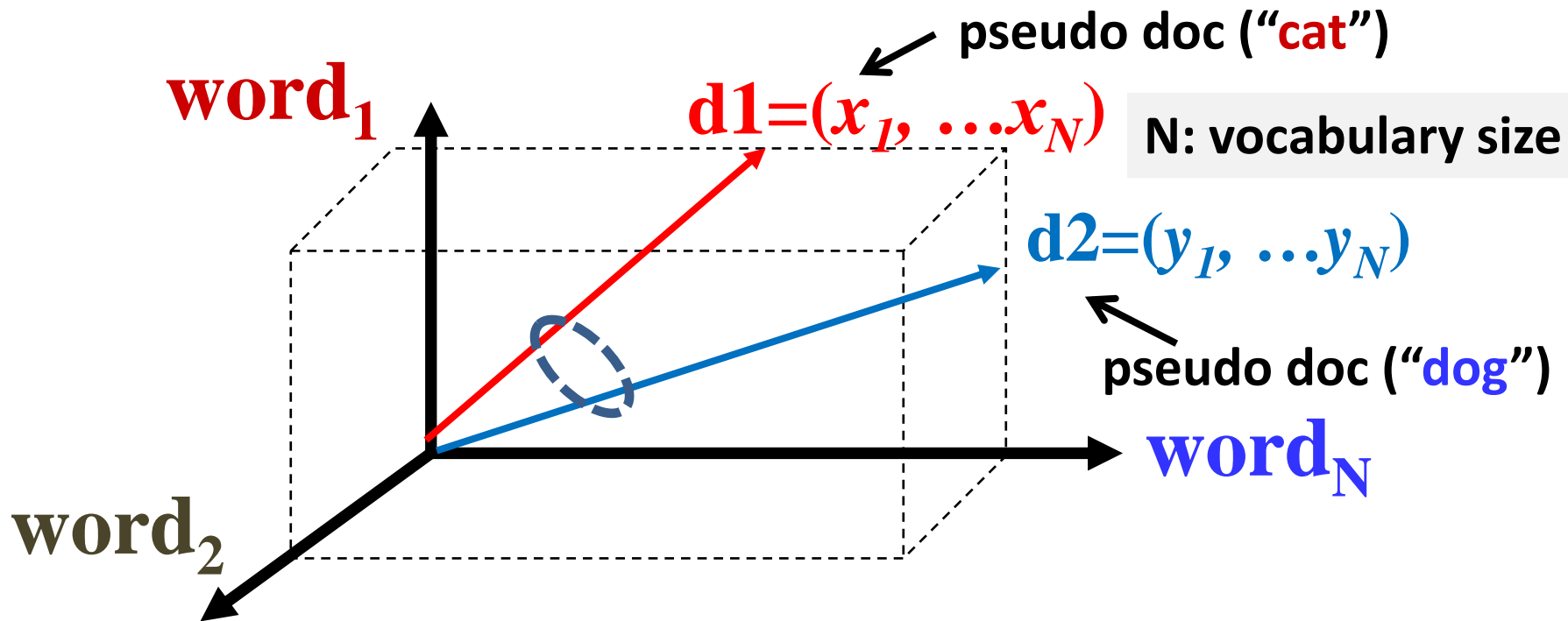
Terms:		"eats"	"ate"	"is"	"has"	….	
Vector:	(	5,	3,	10,	3	….	)