Generalized Vector Space Model (M. Wong, 1985):

Theoretical Claims:

- Defined a better basis for term representation
- Established a way to measure similarities between terms

  i.e., removed the *term independence* assumption from VSM

Example.

matrix A

	d1	d2	d3	q
Monday	1	0	0	0
Tuesday	0	1	1	1
seminar	1	1	0	1
machine	1	1	1	0
translation	1	0	0	0
learning	0	1	0	0
install	0	0	1	0

transformed q and docs

1	2	1	q’
4	2	1	d1’
2	4	2	d2’
1	2	3	d3’

q’ = A^t q

d’ = A^t d

Simplified GVSM: using row vectors without normalization

$$sim(\vec{q}, \vec{d}) = \cos(A^t\vec{q}, A^t\vec{d})$$

- Use term combinations as the basis instead of terms
- Use the *dual space* instead of the original VS
- Use the occurrence pattern of each word across documents
- Allow to measure the similarity between terms
- Transform doc (query) by adding patterns of individual terms