

	トークン1	トークン2	トークン3	トークン4
トークン1				
トークン2				
トークン3				
トークン4				

A

$= \text{Softmax}$

	トークン1	トークン2	トークン3	トークン4
トークン1				
トークン2				
トークン3				
トークン4				

\sqrt{d}

$Q^t K$

トークン1			...	
トークン2			...	
トークン3			...	
トークン4			...	

$=$

	トークン1	トークン2	トークン3	トークン4
トークン1				
トークン2				
トークン3				
トークン4				

A

\times

		...	
		...	
		...	
		...	

V

次の特徴量