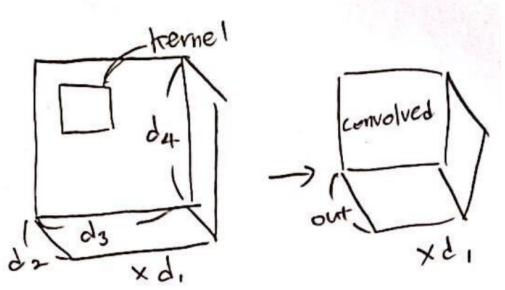
Matmul Layers

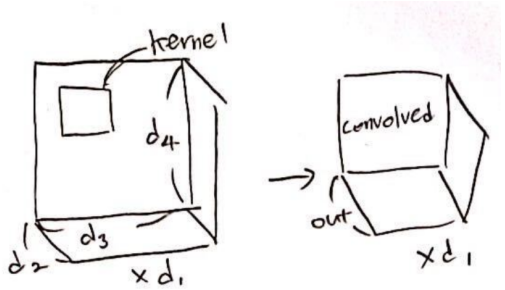
`torch.nn.Linear`

`torch.nn.Linear(in_features, out_features, bias=True)(x)`	
	Require
	• $\|x\| = (d_1, d_2, \ldots, d_k)$ • $\texttt{rank}(\|x\|) \geq 1$ • $d_k = in\_features$
	Guarantees
	• $\|y\| = (d_1, d_2, \ldots, d_{k-1}, out\_features)$
	Comment
	• $y = xA^T + b$를 계산하는 dense 레이어 • 1차원인 경우에도 잘 작동합니다. • *bias* 옵션은 출력 shape에 영향을 주지 않습니다.

$$\frac{\begin{array}{l}\sigma \vdash E \Rightarrow e, c \\ e' = e[1:k-1]@(out) \\ c' = \{(\texttt{rank}(e) \geq 1) \wedge (d_k = in)\}\end{array}}{\sigma \vdash \texttt{Linear}(in, out, bias = True)(E) \Rightarrow e', c \cup c'}$$

Activations

`torch.nn.ReLU`, `torch.nn.ReLU6`, `torch.relu`, `torch.nn.functional.relu`

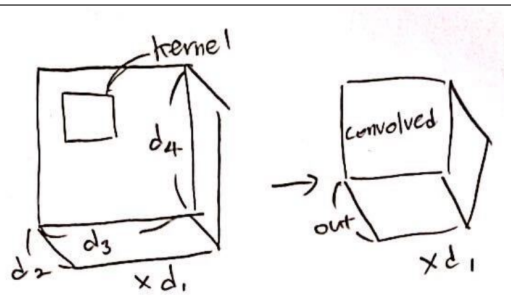
`torch.nn.ReLU(inplace=True)(x)`	
	Require
	Guarantees
	• $\|y\| = \|x\|$ (same shape)
	Comment
	• *inplace* 옵션은 shape에 영향을 주지 않습니다. • `ReLU6`도 `ReLU`와 똑같은 방식으로 shape 계산 • Bulitins인 `torch.relu`와 `torch.nn.functional.relu`는 같은거

$$\forall \texttt{ft} \in \{\texttt{ReLU}, \texttt{ReLU6}\}, \quad \frac{\sigma \vdash E \Rightarrow e, c}{\sigma \vdash \texttt{ft}(inplace = True)(E) \Rightarrow e, c}$$

$$\frac{\sigma \vdash E \Rightarrow e, c}{\sigma \vdash \texttt{relu}(E, inplace = True) \Rightarrow e, c}$$

Technique

`torch.nn.Dropout`, `torch.dropout`, `torch.nn.functional.dropout`

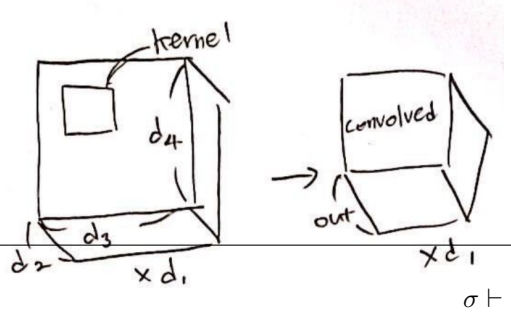
`torch.nn.Dropout(...)(x)`	
	Require
	Guarantees
	• $\|y\| = \|x\|$ (same shape)
	Comment
	• 모든 옵션은 shape에 영향을 주지 않습니다. • Bulitins인 `torch.dropout`와 `torch.nn.functional.dropout`는 서로 역할이 같습니다.

$$\frac{\sigma \vdash E \Rightarrow e, c}{\sigma \vdash \texttt{Dropout}(...)(E) \Rightarrow e, c}$$

$$\frac{\sigma \vdash E \Rightarrow e, c}{\sigma \vdash \texttt{dropout}(E, ...) \Rightarrow e, c}$$

Wrapper

`torch.nn.Sequential`

`torch.nn.Sequential(l1, l2, l3, ..., ln)(x)`	
	Require
	• 순차적으로 shape이 맞아떨어져야함
	Guarantees
	• $\|y\| = \|l_n \circ l_{n-1} \circ \cdots \circ l_1(x)\|$

$$\frac{\sigma \vdash l_n \circ l_{n-1} \circ \cdots l_1(E) \Rightarrow e, c}{\sigma \vdash \texttt{Sequential}(l_1, l_2, \ldots, l_n)(E) \Rightarrow e, c}$$