Activation Functions

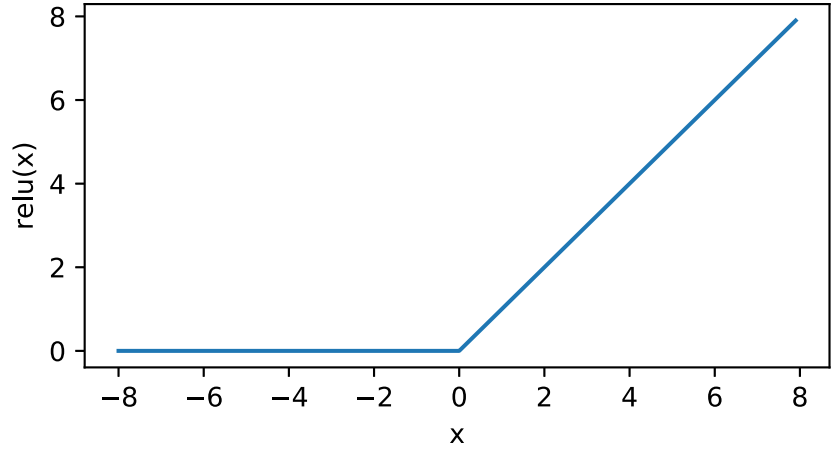
In [1]:
```
%matplotlib inline
import d2l
from mxnet import autograd, nd

def xyplot(x_vals, y_vals, name):
    d2l.set_figsize(figsize=(5, 2.5))
    d2l.plt.plot(x_vals.asnumpy(), y_vals.asnumpy())
    d2l.plt.xlabel('x')
    d2l.plt.ylabel(name + '(x)')
```

ReLU Function

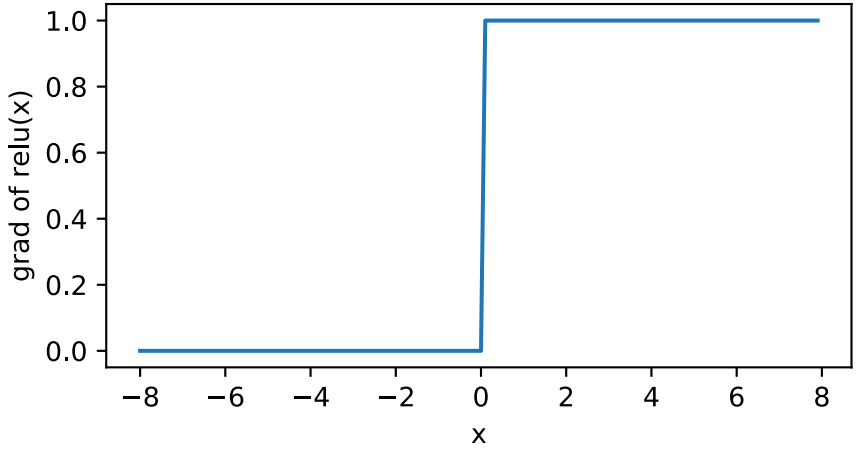
$$\text{ReLU}(x) = \max(x, 0).$$

In [2]:
```
x = nd.arange(-8.0, 8.0, 0.1)
x.attach_grad()
with autograd.record():
    y = x.relu()
xyplot(x, y, 'relu')
```



The Sub-derivative of ReLU

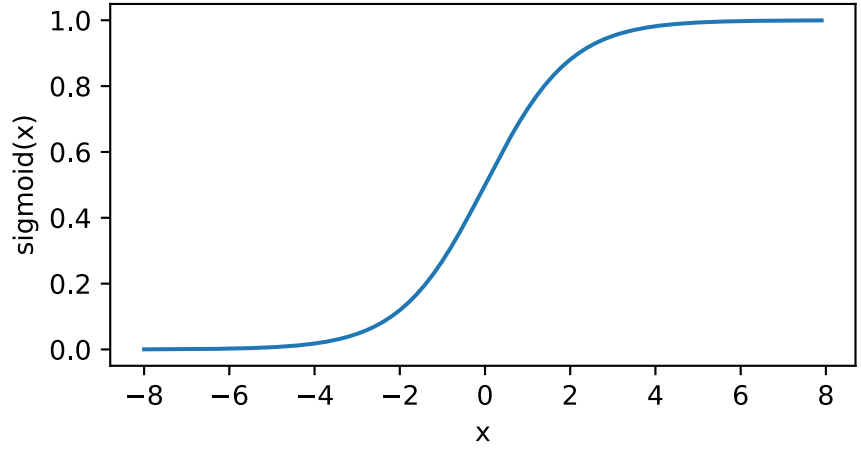
In [3]:
```
y.backward()
xyplot(x, x.grad, 'grad of relu')
```



Sigmoid Function

$$\text{sigmoid}(x) = \frac{1}{1 + \exp(-x)}.$$

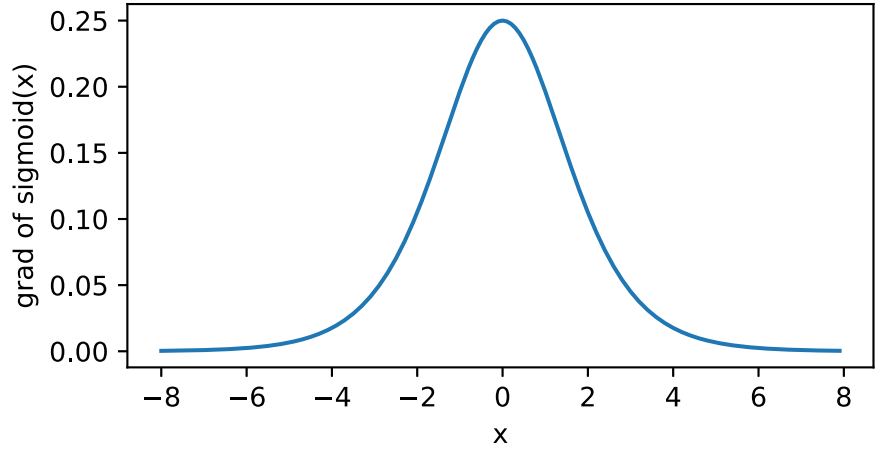
In [4]:
```
with autograd.record():
    y = x.sigmoid()
xyplot(x, y, 'sigmoid')
```



The Derivative of Sigmoid

$$\frac{d}{dx}\text{sigmoid}(x) = \frac{\exp(-x)}{(1+\exp(-x))^2} = \text{sigmoid}(x)\,(1-\text{sigmoid}(x))\,.$$

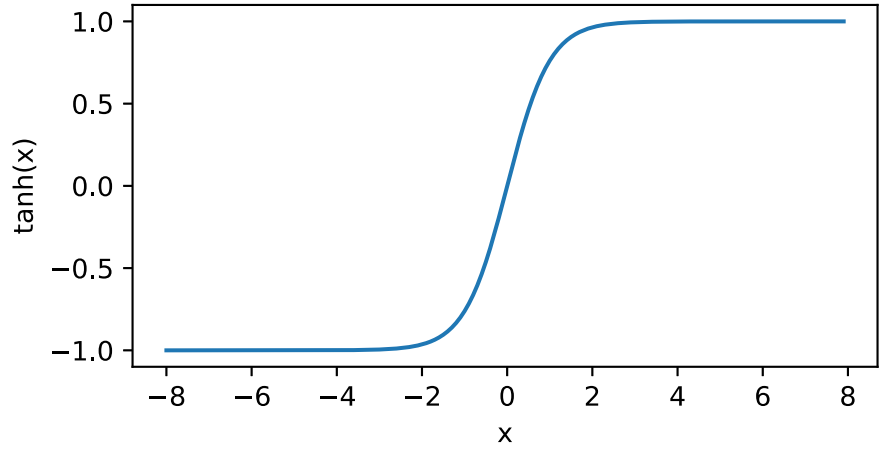
In [5]:
```
y.backward()
xyplot(x, x.grad, 'grad of sigmoid')
```



Tanh Function

$$\tanh(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)}.$$

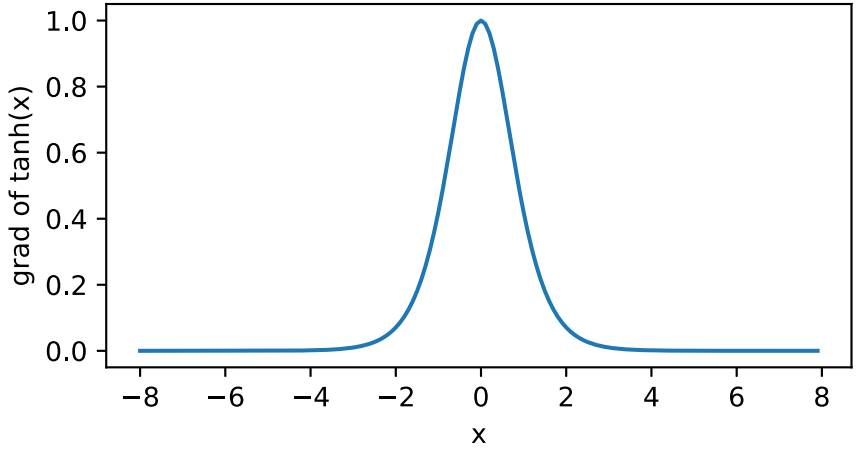
In [6]:
```
with autograd.record():
    y = x.tanh()
xyplot(x, y, 'tanh')
```



The derivative of Tanh

$$\frac{d}{dx}\tanh(x) = 1 - \tanh^2(x).$$

In [7]:
```
y.backward()
xyplot(x, x.grad, 'grad of tanh')
```



In [ ]: