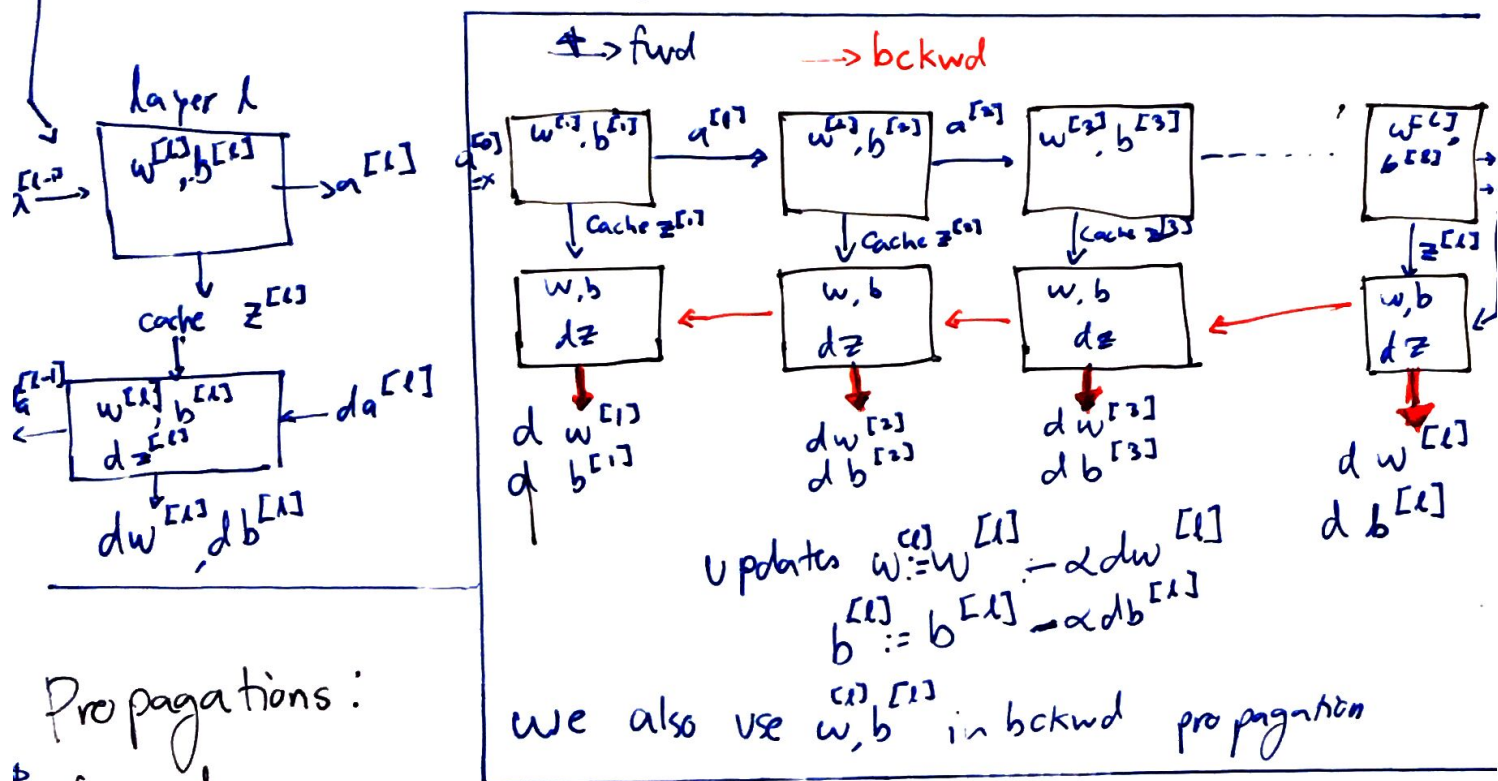# Building blocks of a deep NN

→ visualizing ideas behind fwd, bckwd propagation

for layer l, $w^{[l]}, b^{[l]}$

forward prg^n → $z^{[l]} = w^{[l]} a^{[l-1]} + b^{[l]}$, $a^{[l]} = g^{[l]}(z^{[l]})$
& cache $z^{[l]}$ for bckwd p.

Backward propagation: Input $da^{[l]}$, $\left[\text{output } dg^{[l-1]} \quad dw^{[l]} \quad db^{[l]} \right]$
Cache $z^{[l]}$,

↑ → fwd   → bckwd

layer l

updates $w^{[l]} := w^{[l]} - \alpha \, dw^{[l]}$
$b^{[l]} := b^{[l]} - \alpha \, db^{[l]}$

we also use $w^{[l]}, b^{[l]}$ in bckwd propagation

Propagations:

- forward

inputs $a^{[l-1]}$ outputs $a^{[l]}$ & caches $z^{[l]}$

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$
$$A^{[l]} = g^{[l]}(Z^{[l]})$$

} vectorized version

- backward
inputs $da^{[l]}$, outputs $da^{[l-1]}, dW^{[l]}, db^{[l]}$

$dz^{[l]} = da^{[l]} * g^{[l]\prime}(z^{[l]})$ $* = $ pairwise mult.

$dw^{[l]} = dz^{[l]} * a^{[l-1]T}$

$db^{[l]} = dz^{[l]}$; $da = w^{[l]T} \cdot dz^{[l]}$

$dz^{[l]} = w^{[l+1]T} dz^{[l+1]} * g^{[l]\prime}(z^{[l]})$