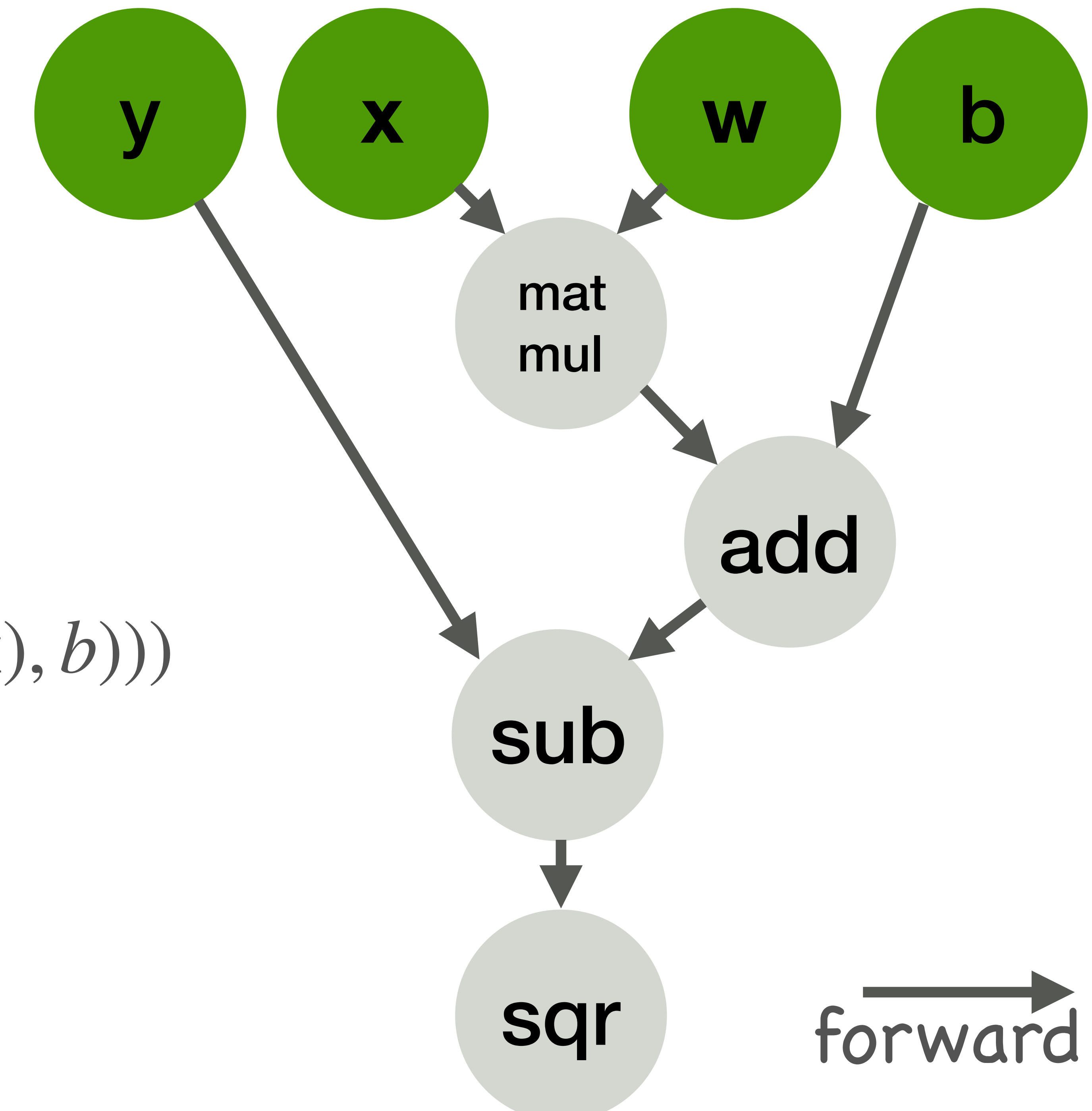
Gradients on computation graphs

Linear regression

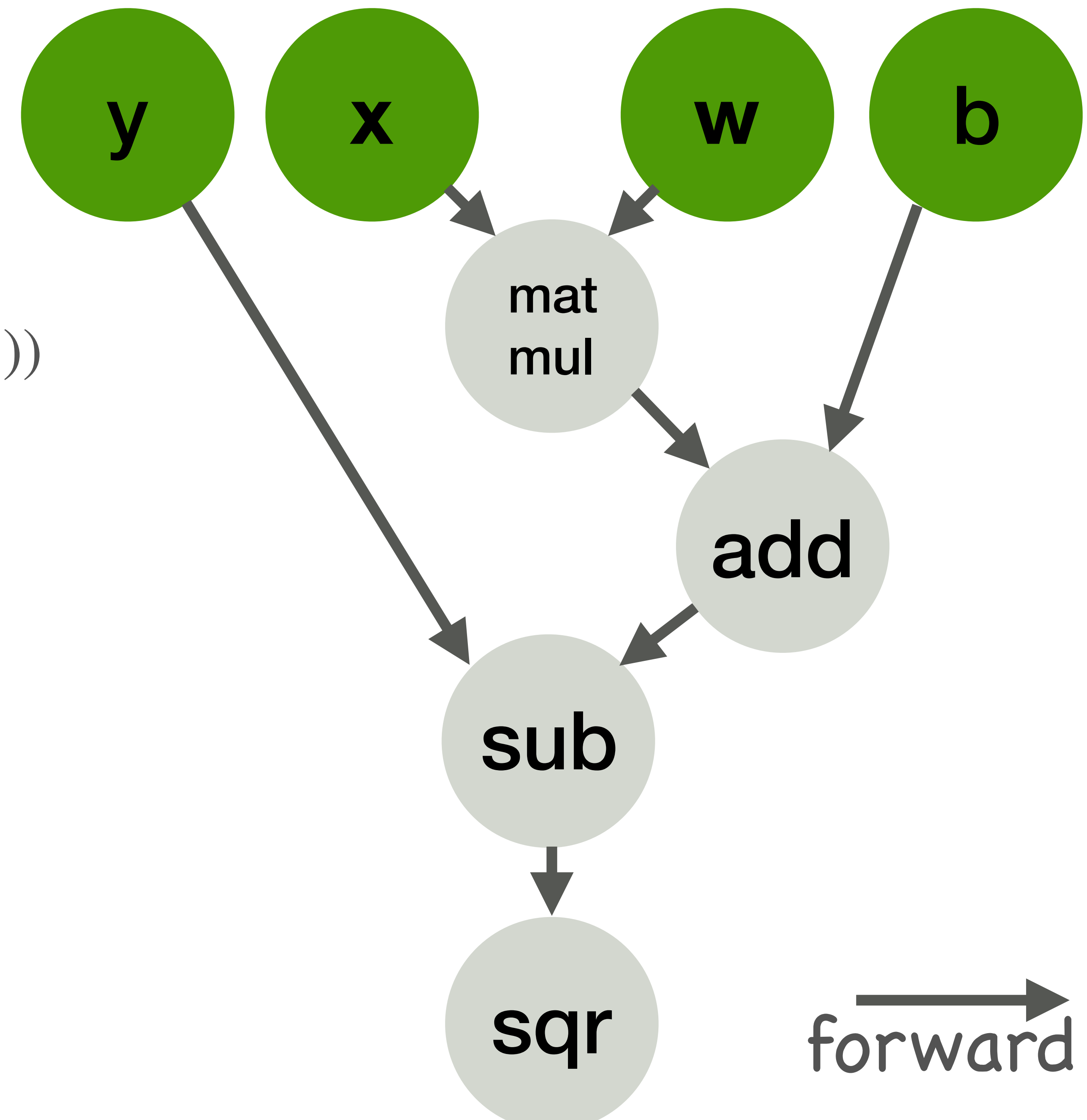
- $\ell(\mathbf{w}, b) = (y - \mathbf{w}^\top \mathbf{x} + b)^2$
- $\mathrm{sqr}(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)))$



Gradient of linear regression

- $$\frac{d\ell(\mathbf{w}, b)}{d\mathbf{w}} = \frac{d(y - \mathbf{w}^\top \mathbf{x} + b)^2}{d\mathbf{w}}$$

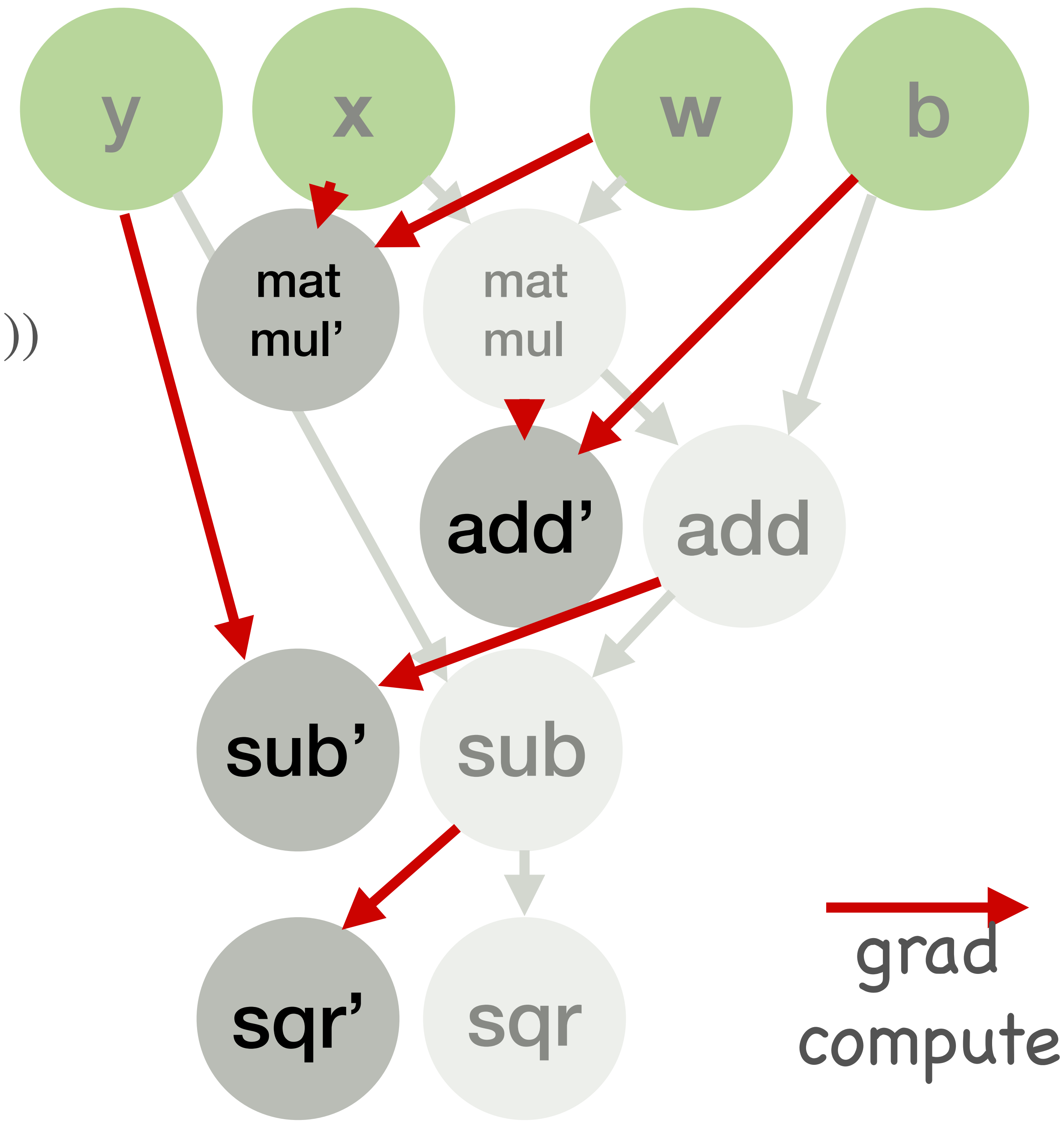
- $$\begin{aligned}
&\frac{d}{d\mathbf{w}}\mathrm{sqr}(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\
&= \mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\
&\quad \frac{d}{d\mathbf{w}}\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)) \\
&= \ldots \\
&= \mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\
&\quad \mathrm{sub}'(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)) \\
&\quad \mathrm{add}'(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b) \\
&\quad \frac{d}{d\mathbf{w}}\mathrm{matmul}(\mathbf{w}, \mathbf{x})
\end{aligned}$$



Gradient of linear regression

- $$\frac{d\ell(\mathbf{w}, b)}{d\mathbf{w}} = \frac{d(y - \mathbf{w}^\top \mathbf{x} + b)^2}{d\mathbf{w}}$$

- $$\begin{aligned} &\frac{d}{d\mathbf{w}}\mathrm{sqr}(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\ &= \mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\ &\quad \frac{d}{d\mathbf{w}}\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)) \\ &= \ldots \\ &= \mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))) \\ &\quad \mathrm{sub}'(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)) \\ &\quad \mathrm{add}'(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b) \\ &\quad \frac{d}{d\mathbf{w}}\mathrm{matmul}(\mathbf{w}, \mathbf{x}) \end{aligned}$$



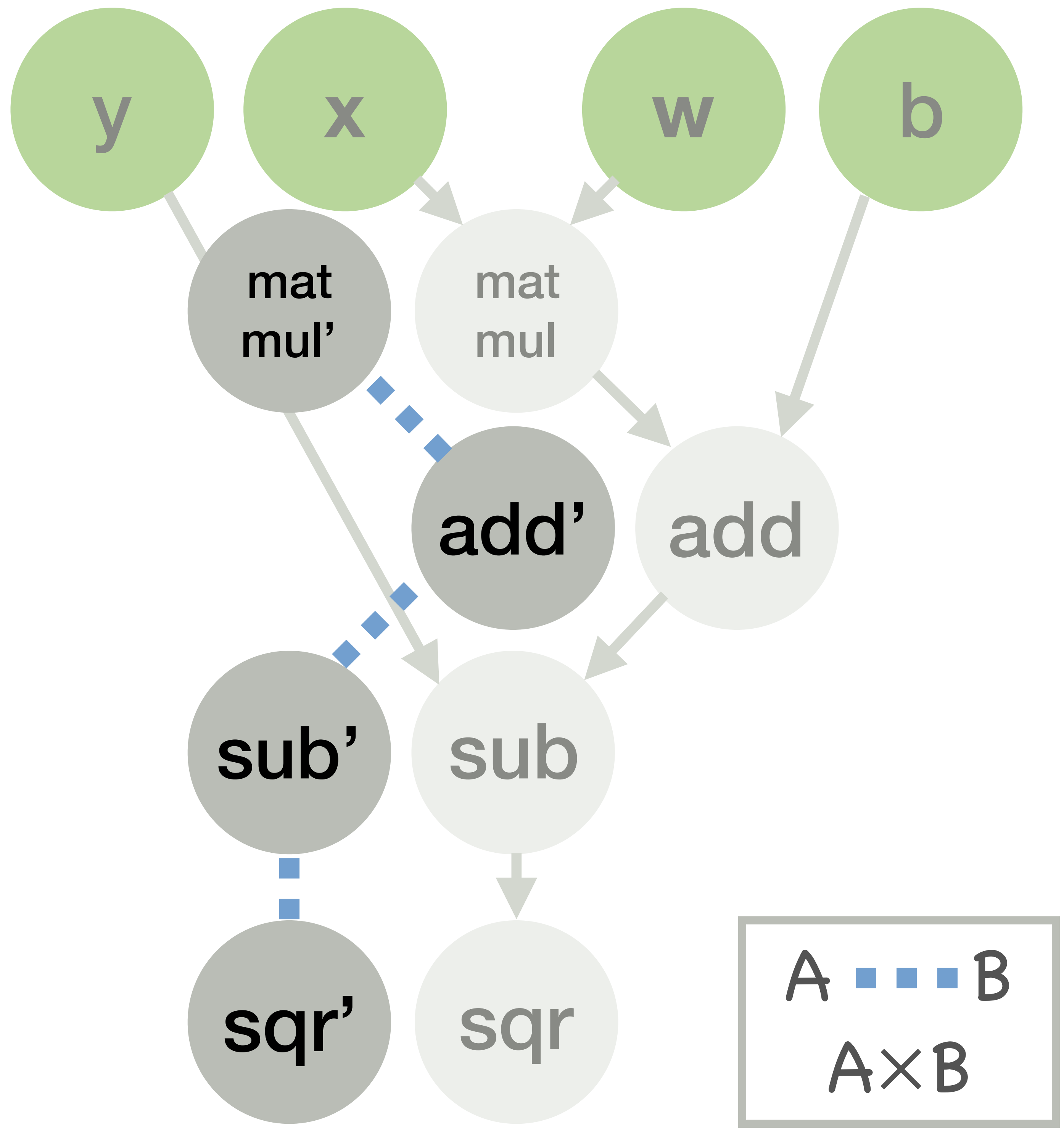
Gradient of linear regression

$\mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)))$

$\mathrm{sub}'(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))$

$\mathrm{add}'(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)$

$$\frac{d}{d\mathbf{w}}\mathrm{matmul}(\mathbf{w}, \mathbf{x})$$



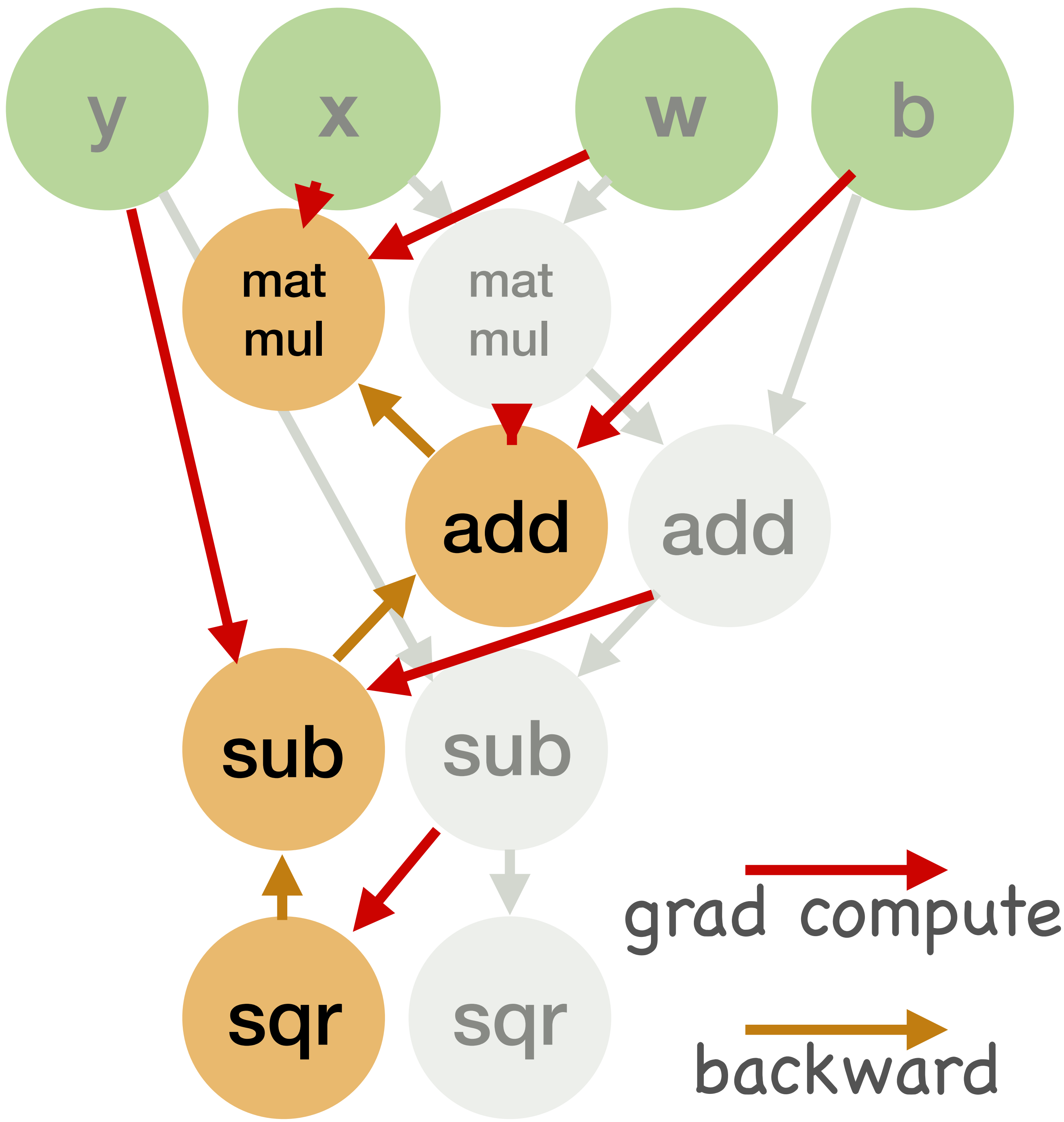
Gradient of linear regression

$\mathrm{sqr}'(\mathrm{sub}(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)))$

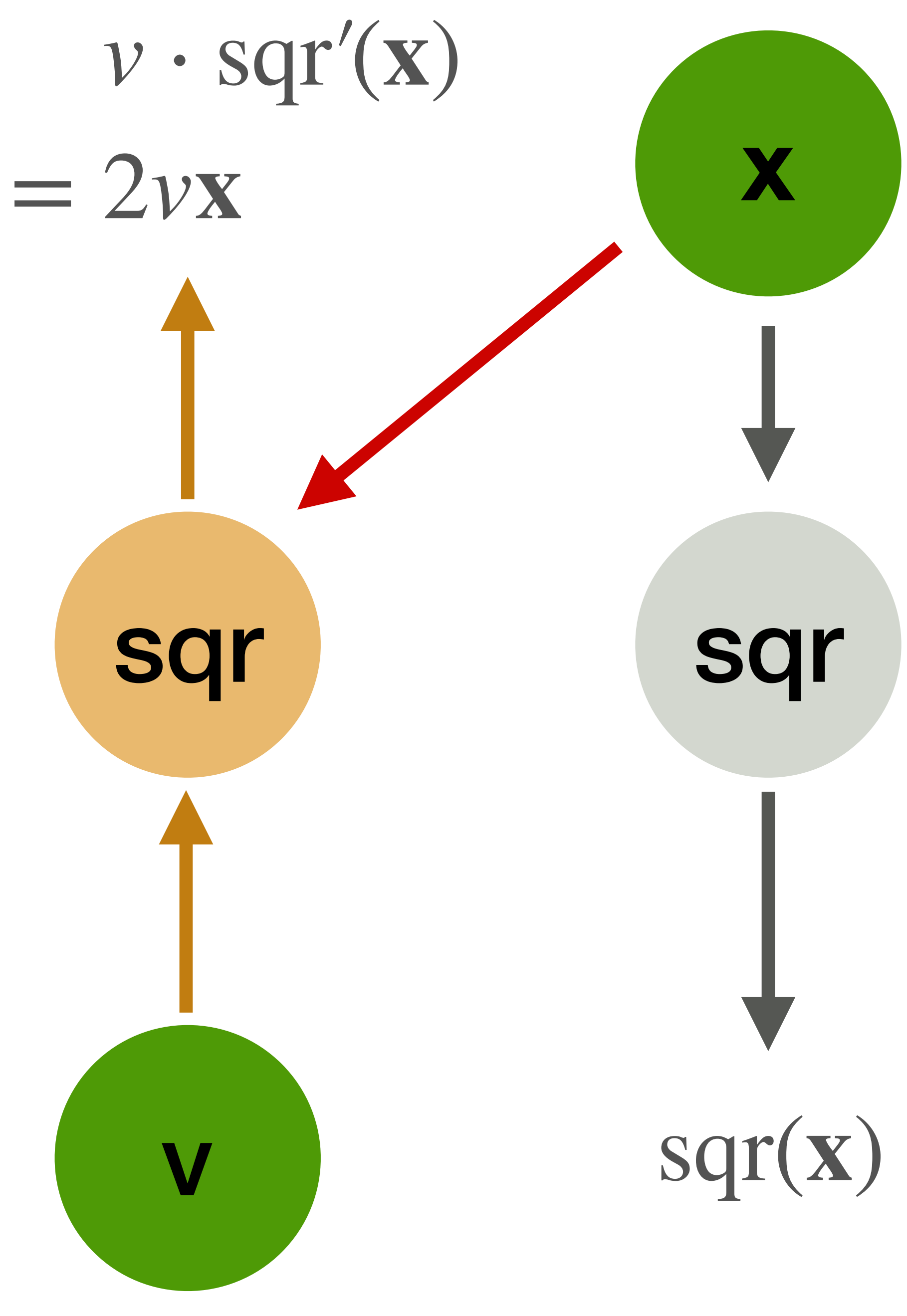
$\mathrm{sub}'(y, \mathrm{add}(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b))$

$\mathrm{add}'(\mathrm{matmul}(\mathbf{w}, \mathbf{x}), b)$

$$\frac{d}{d\mathbf{w}}\mathrm{matmul}(\mathbf{w}, \mathbf{x})$$



Gradient of linear regression



Gradient of linear regression