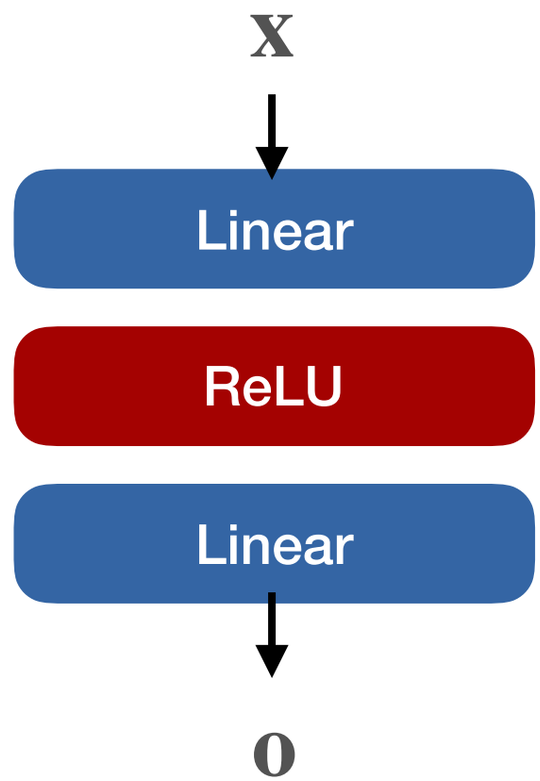
Input normalization

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

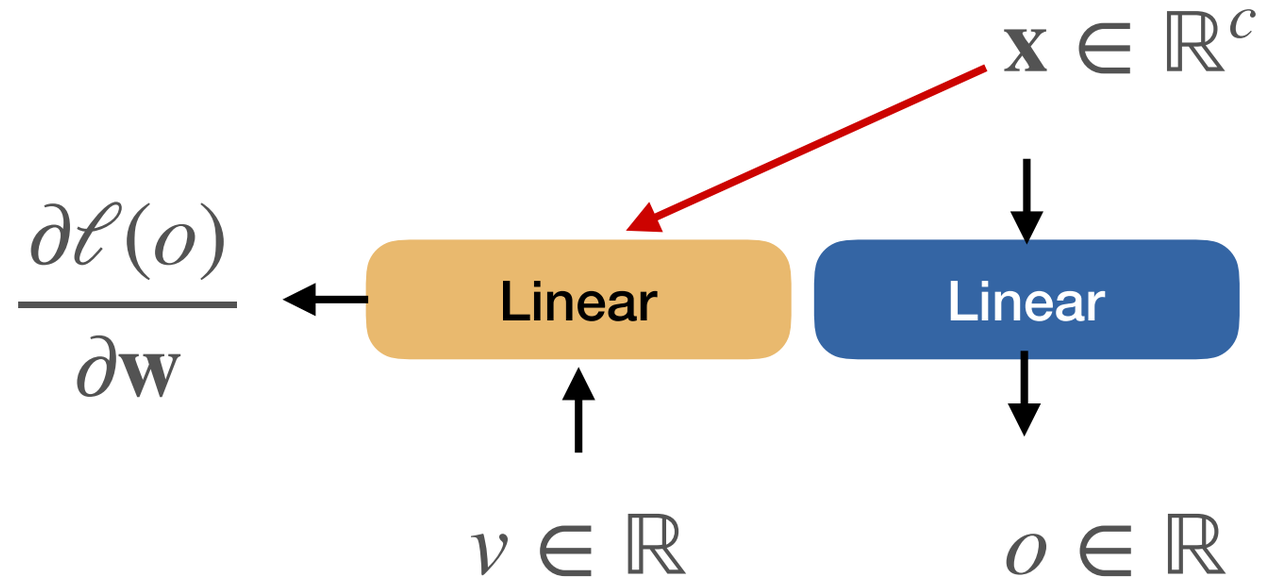
Input normalization

- Input: $\mathbf{x}_i$
- Apply affine transformation $\hat{\mathbf{x}}_i = \alpha \mathbf{x}_i + \beta$



Gradients of uncentered inputs: A simple example

- Input vector $\mathbf{x}$
- Output scalar o
- $\dfrac{\partial \ell(o)}{\partial \mathbf{w}} = v\mathbf{x}^\top$



Mean subtraction

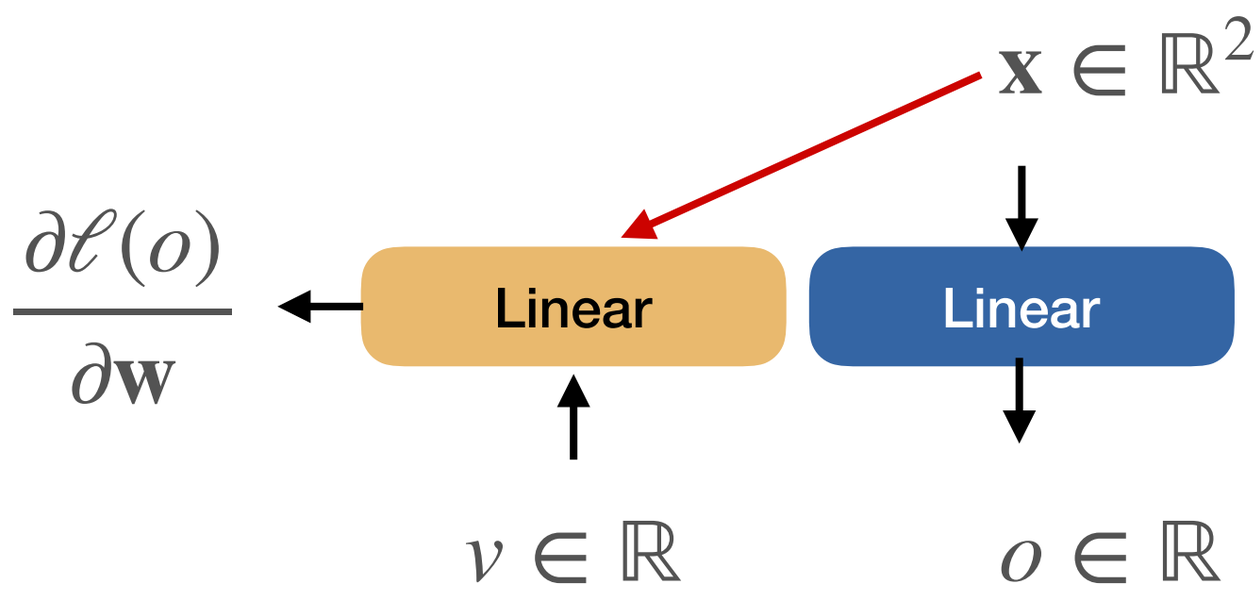
- Input: $\mathbf{x}_i$

- Apply affine transformation $\hat{\mathbf{x}}_i = \mathbf{x}_i - \mu_{\mathbf{x}}$

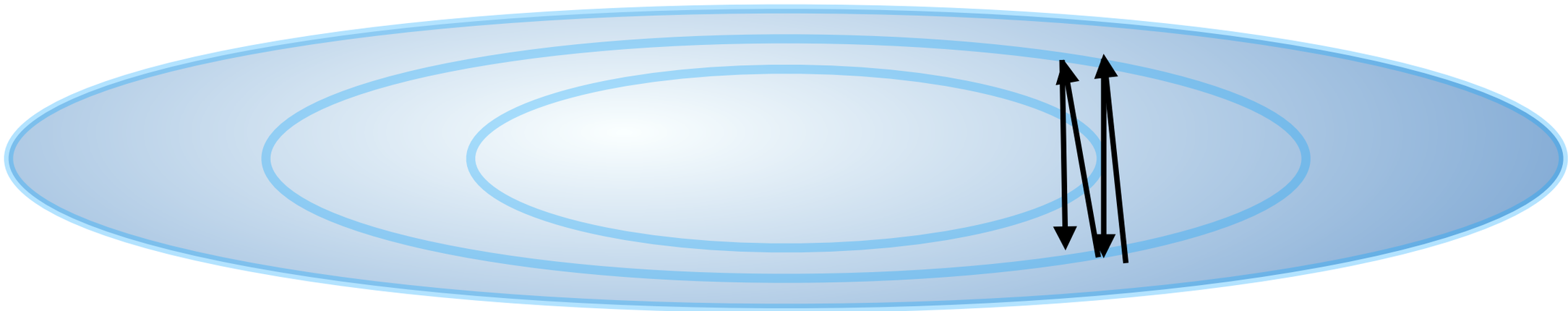
Gradients of unnormalized inputs: A simple example

$$|\mathbf{x}[0]| \ll |\mathbf{x}[1]|$$

- Input vector $\mathbf{x}$
- Output scalar o
- $\frac{\partial \ell(o)}{\partial \mathbf{w}} = v\mathbf{x}^\top$



Gradients of unnormalized inputs: A simple example



Input normalization

- Input: $\mathbf{x}_i$

- Apply affine transformation $\hat{\mathbf{x}}_i = (\mathbf{x}_i - \mu_{\mathbf{x}})/\sigma_{\mathbf{x}}$