5.9 We define

$$g(\mathbf{z}, \boldsymbol{\nu}) := \log\, p(\mathbf{x}, \mathbf{z}) - \log\, q(\mathbf{z}, \boldsymbol{\nu})$$

$$\mathbf{z} := t(\boldsymbol{\epsilon}, \boldsymbol{\nu})$$

for differentiable functions p, q, t, and $\mathbf{x} \in \mathbb{R}^D, \mathbf{z} \in \mathbb{R}^E, \boldsymbol{\nu} \in \mathbb{R}^F, \boldsymbol{\epsilon} \in \mathbb{R}^G$. By using the chain rule, compute the gradient

$$\frac{d}{d\boldsymbol{\nu}} g(\mathbf{z}, \boldsymbol{\nu}).$$

Solution.

Using Chain rule,

$$\frac{d}{d\boldsymbol{\nu}} g(\mathbf{z}, \boldsymbol{\nu}) = \frac{\partial g(\mathbf{z}, \boldsymbol{\nu})}{\partial \mathbf{z}} * \frac{\partial \mathbf{z}}{\partial \boldsymbol{\nu}} + \frac{\partial g(\mathbf{z}, \boldsymbol{\nu})}{\partial \boldsymbol{\nu}}$$

$$= \frac{\partial(\log\, p(\mathbf{x}, \mathbf{z}) - \log\, q(\mathbf{z}, \boldsymbol{\nu}))}{\partial \mathbf{z}} * \frac{\partial t(\boldsymbol{\epsilon}, \boldsymbol{\nu})}{\partial \boldsymbol{\nu}} + \frac{\partial(\log\, p(\mathbf{x}, \mathbf{z}) - \log\, q(\mathbf{z}, \boldsymbol{\nu}))}{\partial \boldsymbol{\nu}}$$

$$= \left(\frac{1}{p(\mathbf{x}, \mathbf{z})} \frac{\partial p(\mathbf{x}, \mathbf{z})}{\partial \mathbf{z}} - \frac{1}{q(\mathbf{z}, \boldsymbol{\nu})} \frac{\partial q(\mathbf{z}, \boldsymbol{\nu})}{\partial \mathbf{z}}\right) \frac{\partial t(\boldsymbol{\epsilon}, \boldsymbol{\nu})}{\partial \boldsymbol{\nu}} + \left(\frac{1}{p(\mathbf{x}, \mathbf{z})} \frac{\partial p(\mathbf{x}, \mathbf{z})}{\partial \boldsymbol{\nu}} - \frac{1}{q(\mathbf{z}, \boldsymbol{\nu})} \frac{\partial q(\mathbf{z}, \boldsymbol{\nu})}{\partial \boldsymbol{\nu}}\right)$$