Derivative explanation

In the video, the derivative of the MSPBE

$$\nabla \left(\mathbf{X}^\top \mathbf{D} \bar{\delta}_{\mathbf{W}}\right)^\top \left(\mathbf{X}^\top \mathbf{D} \mathbf{X}\right)^{-1} \left(\mathbf{X}^\top \mathbf{D} \bar{\delta}_{\mathbf{w}}\right)$$
$$= 2 \left(\nabla \mathbf{X}^\top \mathbf{D} \bar{\delta}_{\mathbf{W}}\right)^\top \left(\mathbf{X}^\top \mathbf{D} \mathbf{X}\right)^{-1} \left(\mathbf{X}^\top \mathbf{D} \bar{\delta}_{\mathbf{w}}\right)$$

is compared to the scalar case, where we have

$$\nabla f(\boldsymbol{w}) c f(\boldsymbol{w}) = 2(\nabla f(\boldsymbol{w})) c f(\boldsymbol{w})$$

which is easily verified by applying the chain rule

$$\nabla c f(\boldsymbol{w})^2 = c 2 f(\boldsymbol{w}) \nabla f(\boldsymbol{w})$$

(the 'matrix cookbook' gives a thorough explanation, https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf)