AML5202 | Deep Learning | Problem Set-1

Suppose we have a 3×5-weights matrix $\mathbf{W}$ and a 5-vector $\mathbf{x}$. We want to calculate the gradient $\nabla_{\mathbf{W}}(\mathbf{Wx})$. To that end, answer the following questions:

1. What is the shape of $\mathbf{Wx}$.
2. What is the shape of the gradient $\nabla_{\mathbf{W}}(\mathbf{Wx})$?
3. Suppose the weights matrix is represented as $\mathbf{W} = \begin{bmatrix} \mathbf{w}_1^{\mathrm{T}} \\ \mathbf{w}_2^{\mathrm{T}} \\ \mathbf{w}_3^{\mathrm{T}} \end{bmatrix}$. What is the shape of $\mathbf{w}_1$?
4. What is the shape of gradient $\nabla_{\mathbf{w}_1}(\mathbf{Wx})$?
5. Calculate $\nabla_{\mathbf{w}_1}(\mathbf{Wx})$ using the fact that $\mathbf{Wx} = \begin{bmatrix} \mathbf{w}_1^{\mathrm{T}} \\ \mathbf{w}_2^{\mathrm{T}} \\ \mathbf{w}_3^{\mathrm{T}} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{w}_1^{\mathrm{T}}\mathbf{x} \\ \mathbf{w}_2^{\mathrm{T}}\mathbf{x} \\ \mathbf{w}_3^{\mathrm{T}}\mathbf{x} \end{bmatrix}$.
6. Similarly, calculate $\nabla_{\mathbf{w}_2}(\mathbf{Wx})$ and $\nabla_{\mathbf{w}_3}(\mathbf{Wx})$.
7. Using the results from above, clearly show the entries of the gradient $\nabla_{\mathbf{W}}(\mathbf{Wx})$ as a 3D-tensor.