EN2550 Fundamentals of Image Processing and Machine Vision: Backpropagation for a Liner Network

Ranga Rodrigo
ranga@uom.lk

The University of Moratuwa, Sri Lanka

May 17, 2022

Electronic and Telecommunication Engineering

Backpropagation for a Linear Network

No. of training examples = 5.
No. of features in each training ex. = 4.
No. of classes = 3, one-hot encoding.

$$\boldsymbol{x} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ x_{51} & x_{52} & x_{53} & x_{54} \end{bmatrix}_{5\times 4}$$

$$\hat{\boldsymbol{y}} = \begin{bmatrix} \hat{y}_{11} & \hat{y}_{12} & \hat{y}_{13} \\ \hat{y}_{21} & \hat{y}_{22} & \hat{y}_{23} \\ \hat{y}_{31} & \hat{y}_{32} & \hat{y}_{33} \\ \hat{y}_{41} & \hat{y}_{42} & \hat{y}_{43} \\ \hat{y}_{51} & \hat{y}_{52} & \hat{y}_{53} \end{bmatrix}_{5\times 3}$$

$$\boldsymbol{w} = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix}_{4\times 3}$$

$$\boldsymbol{b} = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}_{1\times 3}$$

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \\ x_{51} & x_{52} & x_{53} & x_{54} \end{bmatrix} \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ w_{31} & w_{32} & w_{33} \\ w_{41} & w_{42} & w_{43} \end{bmatrix} + \boldsymbol{b} = \begin{bmatrix} \hat{y}_{11} & \hat{y}_{12} & \hat{y}_{13} \\ \hat{y}_{21} & \hat{y}_{22} & \hat{y}_{23} \\ \hat{y}_{31} & \hat{y}_{32} & \hat{y}_{33} \\ \hat{y}_{41} & \hat{y}_{42} & \hat{y}_{43} \\ \hat{y}_{51} & \hat{y}_{52} & \hat{y}_{53} \end{bmatrix}$$

$$\hat{\boldsymbol{y}} = \boldsymbol{x}\boldsymbol{w} + \boldsymbol{b}$$

$$L = \frac{1}{5} \sum \begin{bmatrix} (\hat{y}_{11} - y_{11})^2 & (\hat{y}_{12} - y_{12})^2 & (\hat{y}_{13} - y_{13})^2 \\ (\hat{y}_{21} - y_{21})^2 & (\hat{y}_{22} - y_{22})^2 & (\hat{y}_{23} - y_{23})^2 \\ (\hat{y}_{31} - y_{31})^2 & (\hat{y}_{32} - y_{32})^2 & (\hat{y}_{33} - y_{33})^2 \\ (\hat{y}_{41} - y_{41})^2 & (\hat{y}_{42} - y_{42})^2 & (\hat{y}_{43} - y_{43})^2 \\ (\hat{y}_{51} - y_{51})^2 & (\hat{y}_{52} - y_{52})^2 & (\hat{y}_{53} - y_{53})^2 \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial \hat{y}_{11}}\frac{\partial \hat{y}_{11}}{\partial w_{11}} + \frac{\partial L}{\partial \hat{y}_{21}}\frac{\partial \hat{y}_{21}}{\partial w_{11}} + \frac{\partial L}{\partial \hat{y}_{31}}\frac{\partial \hat{y}_{31}}{\partial w_{11}} + \frac{\partial L}{\partial \hat{y}_{41}}\frac{\partial \hat{y}_{41}}{\partial w_{11}} + \frac{\partial L}{\partial \hat{y}_{51}}\frac{\partial \hat{y}_{51}}{\partial w_{11}}$$

$$\frac{\partial L}{\partial w_{11}} = \frac{\partial L}{\partial \hat{y}_{11}}x_{11} + \frac{\partial L}{\partial \hat{y}_{21}}x_{21} + \frac{\partial L}{\partial \hat{y}_{31}}x_{31} + \frac{\partial L}{\partial \hat{y}_{41}}x_{41} + \frac{\partial L}{\partial \hat{y}_{51}}x_{51}$$

$$\frac{\partial L}{\partial w_{11}} = \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{11}} & \frac{\partial L}{\partial \hat{y}_{21}} & \frac{\partial L}{\partial \hat{y}_{31}} & \frac{\partial L}{\partial \hat{y}_{41}} & \frac{\partial L}{\partial \hat{y}_{51}} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \\ x_{51} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{12}} = \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{12}} & \frac{\partial L}{\partial \hat{y}_{22}} & \frac{\partial L}{\partial \hat{y}_{32}} & \frac{\partial L}{\partial \hat{y}_{42}} & \frac{\partial L}{\partial \hat{y}_{52}} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \\ x_{51} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{13}} = \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{13}} & \frac{\partial L}{\partial \hat{y}_{23}} & \frac{\partial L}{\partial \hat{y}_{33}} & \frac{\partial L}{\partial \hat{y}_{43}} & \frac{\partial L}{\partial \hat{y}_{53}} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ x_{41} \\ x_{51} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_{21}} = \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{11}} & \frac{\partial L}{\partial \hat{y}_{21}} & \frac{\partial L}{\partial \hat{y}_{31}} & \frac{\partial L}{\partial \hat{y}_{41}} & \frac{\partial L}{\partial \hat{y}_{51}} \end{bmatrix} \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \\ x_{42} \\ x_{52} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial L}{\partial w_{11}} & \frac{\partial L}{\partial w_{12}} & \frac{\partial L}{\partial w_{13}} \\ \frac{\partial L}{\partial w_{21}} & \frac{\partial L}{\partial w_{22}} & \frac{\partial L}{\partial w_{23}} \\ \frac{\partial L}{\partial w_{31}} & \frac{\partial L}{\partial w_{32}} & \frac{\partial L}{\partial w_{33}} \\ \frac{\partial L}{\partial w_{41}} & \frac{\partial L}{\partial w_{42}} & \frac{\partial L}{\partial w_{43}} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & x_{31} & x_{41} & x_{51} \\ x_{12} & x_{22} & x_{32} & x_{42} & x_{52} \\ x_{13} & x_{23} & x_{33} & x_{43} & x_{53} \\ x_{14} & x_{24} & x_{34} & x_{44} & x_{54} \end{bmatrix} \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{11}} & \frac{\partial L}{\partial \hat{y}_{12}} & \frac{\partial L}{\partial \hat{y}_{13}} \\ \frac{\partial L}{\partial \hat{y}_{21}} & \frac{\partial L}{\partial \hat{y}_{22}} & \frac{\partial L}{\partial \hat{y}_{23}} \\ \frac{\partial L}{\partial \hat{y}_{31}} & \frac{\partial L}{\partial \hat{y}_{32}} & \frac{\partial L}{\partial \hat{y}_{33}} \\ \frac{\partial L}{\partial \hat{y}_{41}} & \frac{\partial L}{\partial \hat{y}_{42}} & \frac{\partial L}{\partial \hat{y}_{43}} \\ \frac{\partial L}{\partial \hat{y}_{51}} & \frac{\partial L}{\partial \hat{y}_{52}} & \frac{\partial L}{\partial \hat{y}_{53}} \end{bmatrix}$$

$$\frac{\partial L}{\partial \boldsymbol{w}} = \boldsymbol{x}^\top \frac{\partial L}{\partial \hat{\boldsymbol{y}}}$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial \hat{y}_{11}}\frac{\partial \hat{y}_{11}}{\partial b_1} + \frac{\partial L}{\partial \hat{y}_{21}}\frac{\partial \hat{y}_{21}}{\partial b_1} + \frac{\partial L}{\partial \hat{y}_{31}}\frac{\partial \hat{y}_{31}}{\partial b_1} + \frac{\partial L}{\partial \hat{y}_{41}}\frac{\partial \hat{y}_{41}}{\partial b_1} + \frac{\partial L}{\partial \hat{y}_{51}}\frac{\partial \hat{y}_{51}}{\partial b_1}$$

$$= \frac{\partial L}{\partial \hat{y}_{11}} \times 1 + \frac{\partial L}{\partial \hat{y}_{21}} \times 1 + \frac{\partial L}{\partial \hat{y}_{31}} \times 1 + \frac{\partial L}{\partial \hat{y}_{41}} \times 1 + \frac{\partial L}{\partial \hat{y}_{51}} \times 1$$

$$\frac{\partial L}{\partial \boldsymbol{b}} = \begin{bmatrix} \frac{\partial L}{\partial b_1} & \frac{\partial L}{\partial b_2} & \frac{\partial L}{\partial b_3} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial L}{\partial \hat{y}_{11}} + \frac{\partial L}{\partial \hat{y}_{21}} + \frac{\partial L}{\partial \hat{y}_{31}} + \frac{\partial L}{\partial \hat{y}_{41}} + \frac{\partial L}{\partial \hat{y}_{51}} \\ \frac{\partial L}{\partial \hat{y}_{12}} + \frac{\partial L}{\partial \hat{y}_{22}} + \frac{\partial L}{\partial \hat{y}_{32}} + \frac{\partial L}{\partial \hat{y}_{42}} + \frac{\partial L}{\partial \hat{y}_{52}} \\ \frac{\partial L}{\partial \hat{y}_{13}} + \frac{\partial L}{\partial \hat{y}_{23}} + \frac{\partial L}{\partial \hat{y}_{33}} + \frac{\partial L}{\partial \hat{y}_{43}} + \frac{\partial L}{\partial \hat{y}_{53}} \end{bmatrix}^{\mathsf{T}}$$

$$\frac{\partial L}{\partial \boldsymbol{b}} = \sum_{\text{columns}} \begin{bmatrix} \frac{\partial L}{\partial \hat{\boldsymbol{y}}} \end{bmatrix}$$