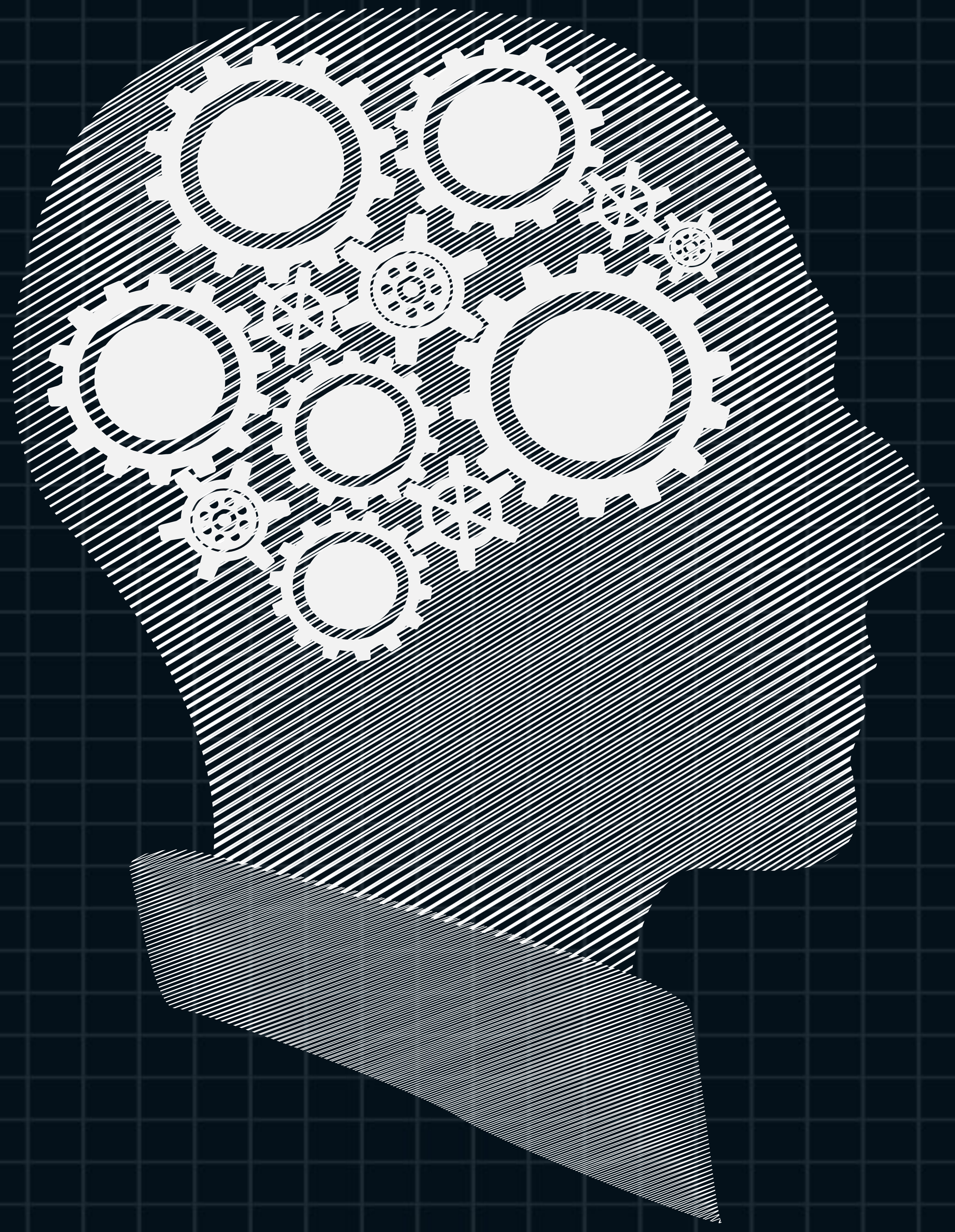
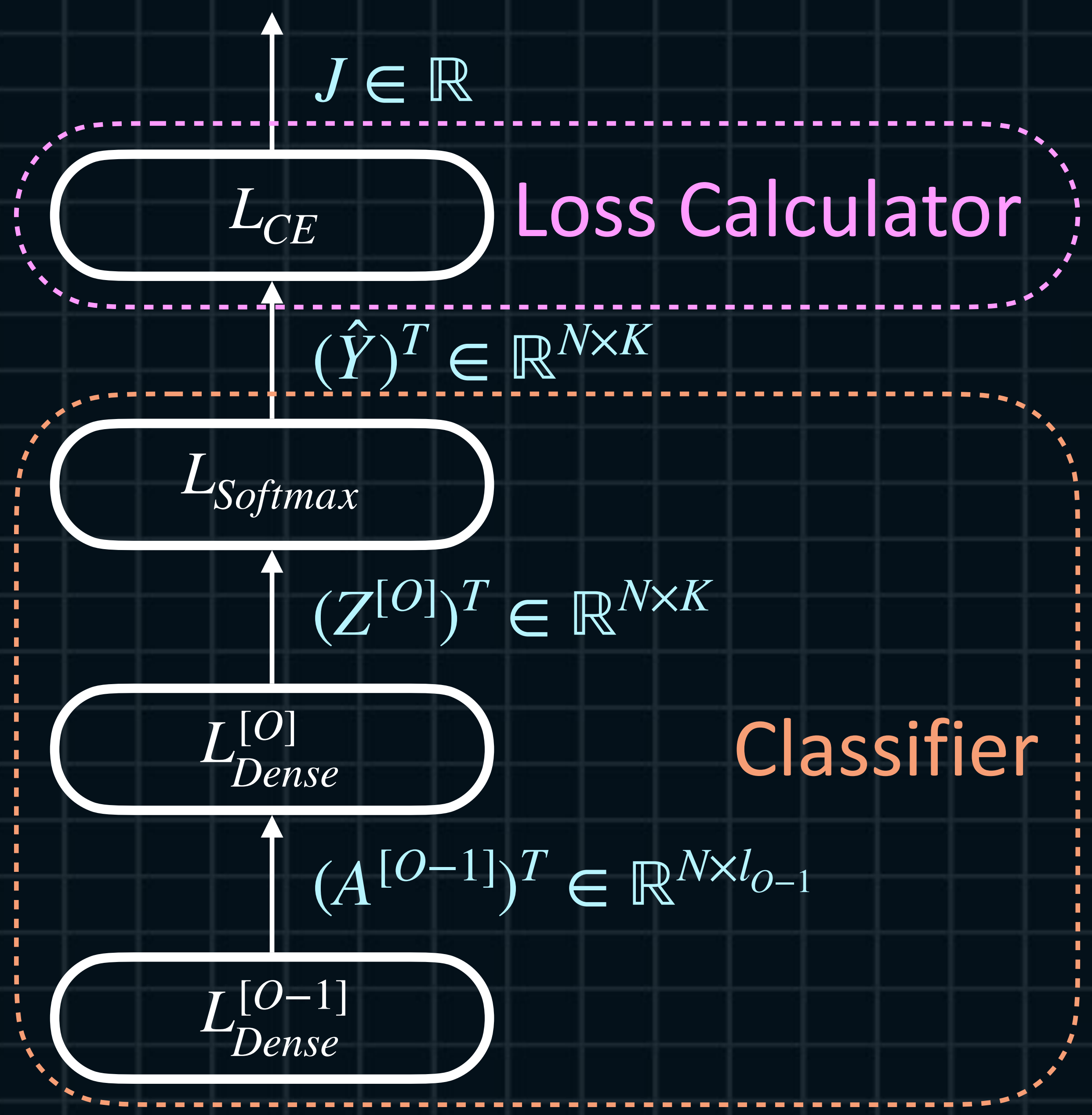


Backpropagation and Jacobian Matrices

Lecture.9 Expansion of Jacobians

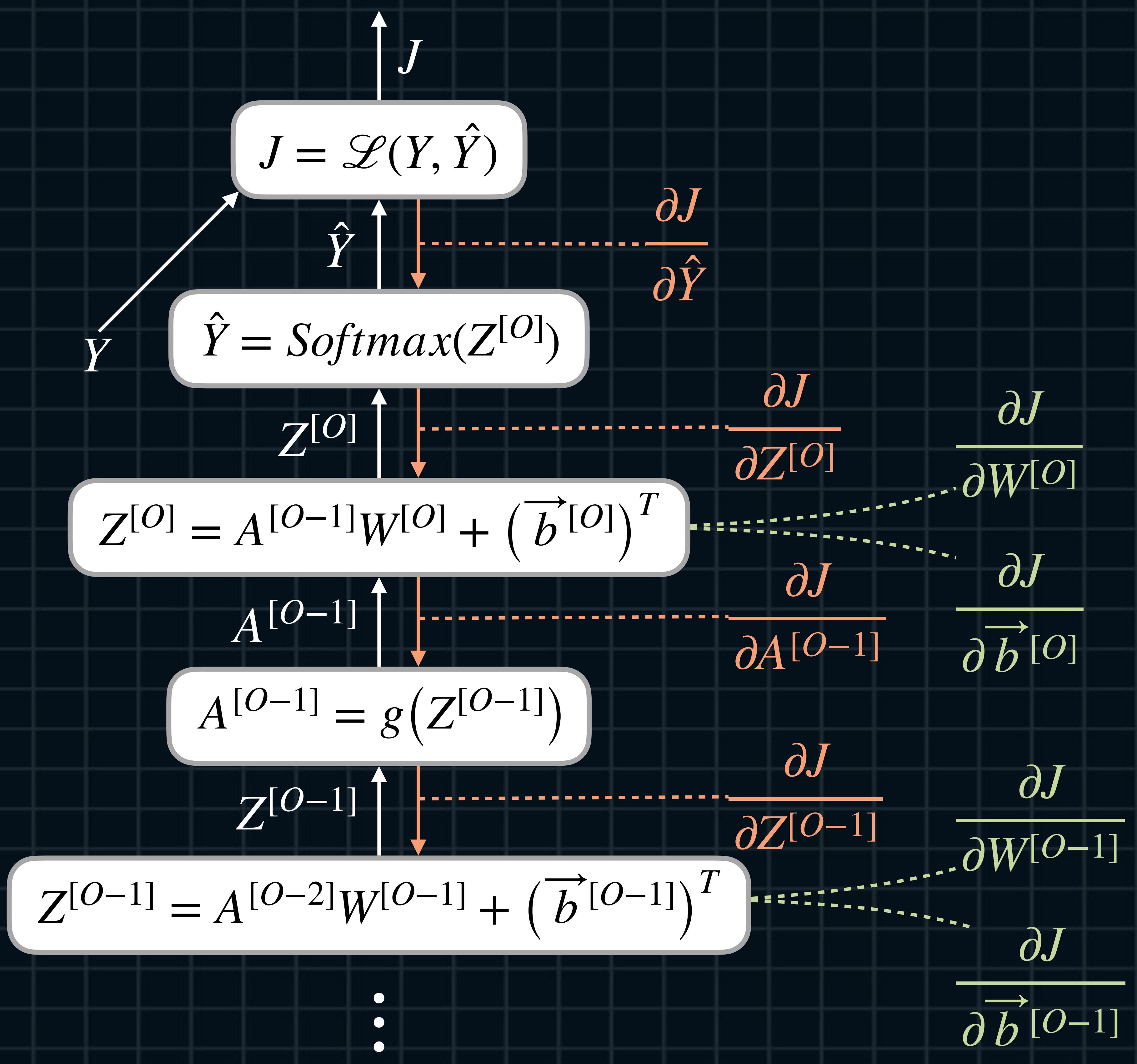
- Tensor Operations and Jacobians

MLP Model and Tensors



Lecture.9 Expansion of Jacobians - Tensor Operations and Jacobians

Jacobians with Matrices



$$\frac{\partial J}{\partial W}, \frac{\partial J}{\partial \vec{b}}$$

- New Notations

Jacobians Matrices

$$M \in \mathbb{R}^{\alpha\times\beta} = (m_{ij}) = \begin{pmatrix} m_{11} & m_{12} & \dots & m_{1\beta} \\ m_{21} & m_{22} & \dots & m_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ m_{\alpha 1} & m_{\alpha 2} & \dots & m_{\alpha\beta} \end{pmatrix}$$

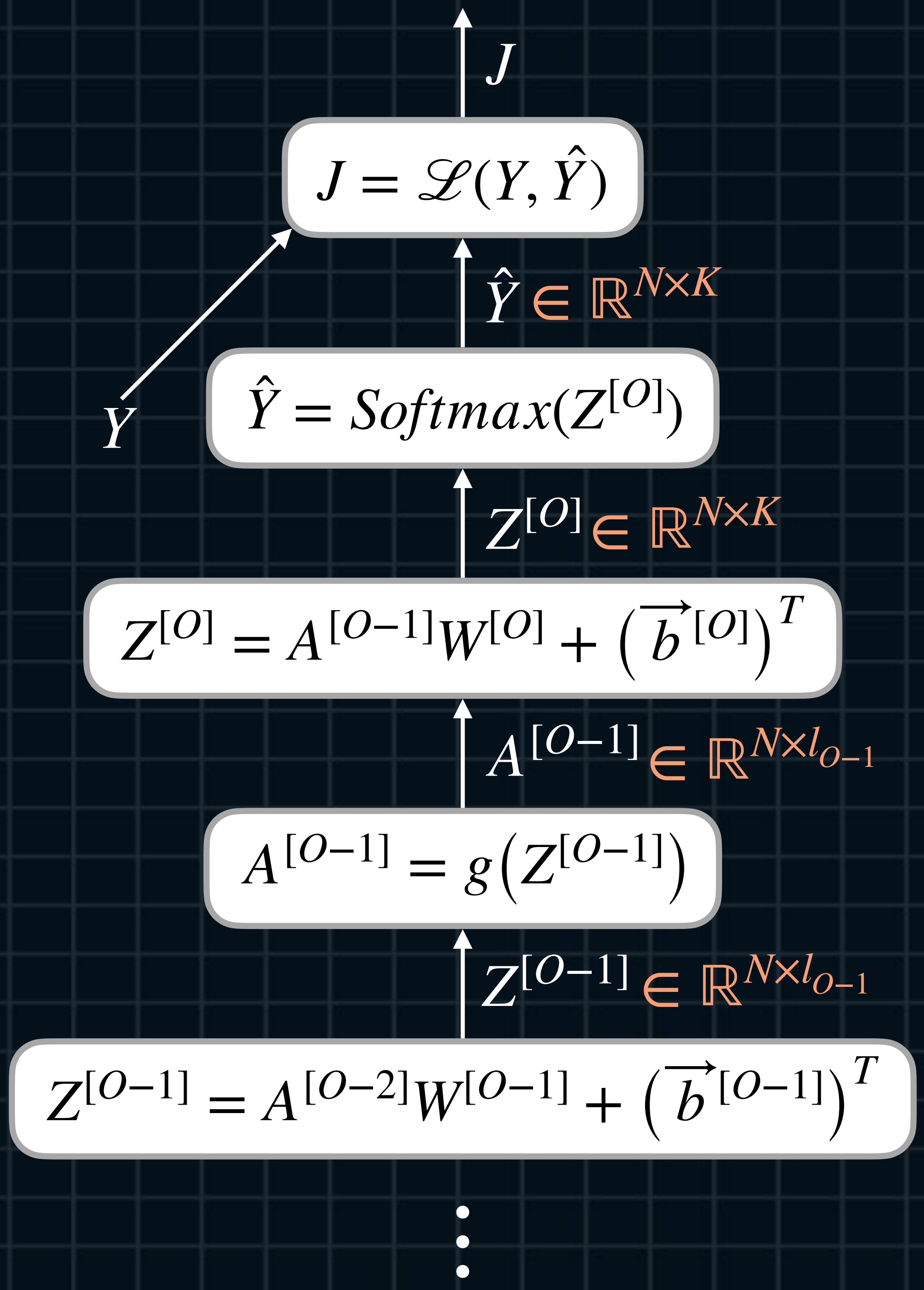
$$\frac{\partial J}{\partial M} = dM, \;\; dm_{ij} = \frac{\partial J}{\partial m_{ij}}$$

$$M \in \mathbb{R}^{\alpha\times\beta} \implies dM \in \mathbb{R}^{\alpha\times\beta}$$

$$dM = \frac{\partial J}{\partial M} = (dm_{ij}) = \begin{pmatrix} dm_{11} & dm_{12} & \dots & dm_{1\beta} \\ dm_{21} & dm_{22} & \dots & dm_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dm_{\alpha 1} & dm_{\alpha 2} & \dots & dm_{\alpha\beta} \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial m_{11}} & \frac{\partial J}{\partial m_{12}} & \dots & \frac{\partial J}{\partial m_{1\beta}} \\ \frac{\partial J}{\partial m_{21}} & \frac{\partial J}{\partial m_{22}} & \dots & \frac{\partial J}{\partial m_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial m_{\alpha 1}} & \frac{\partial J}{\partial m_{\alpha 2}} & \dots & \frac{\partial J}{\partial m_{\alpha\beta}} \end{pmatrix}$$

- New Notations

Jacobians within Neural Networks

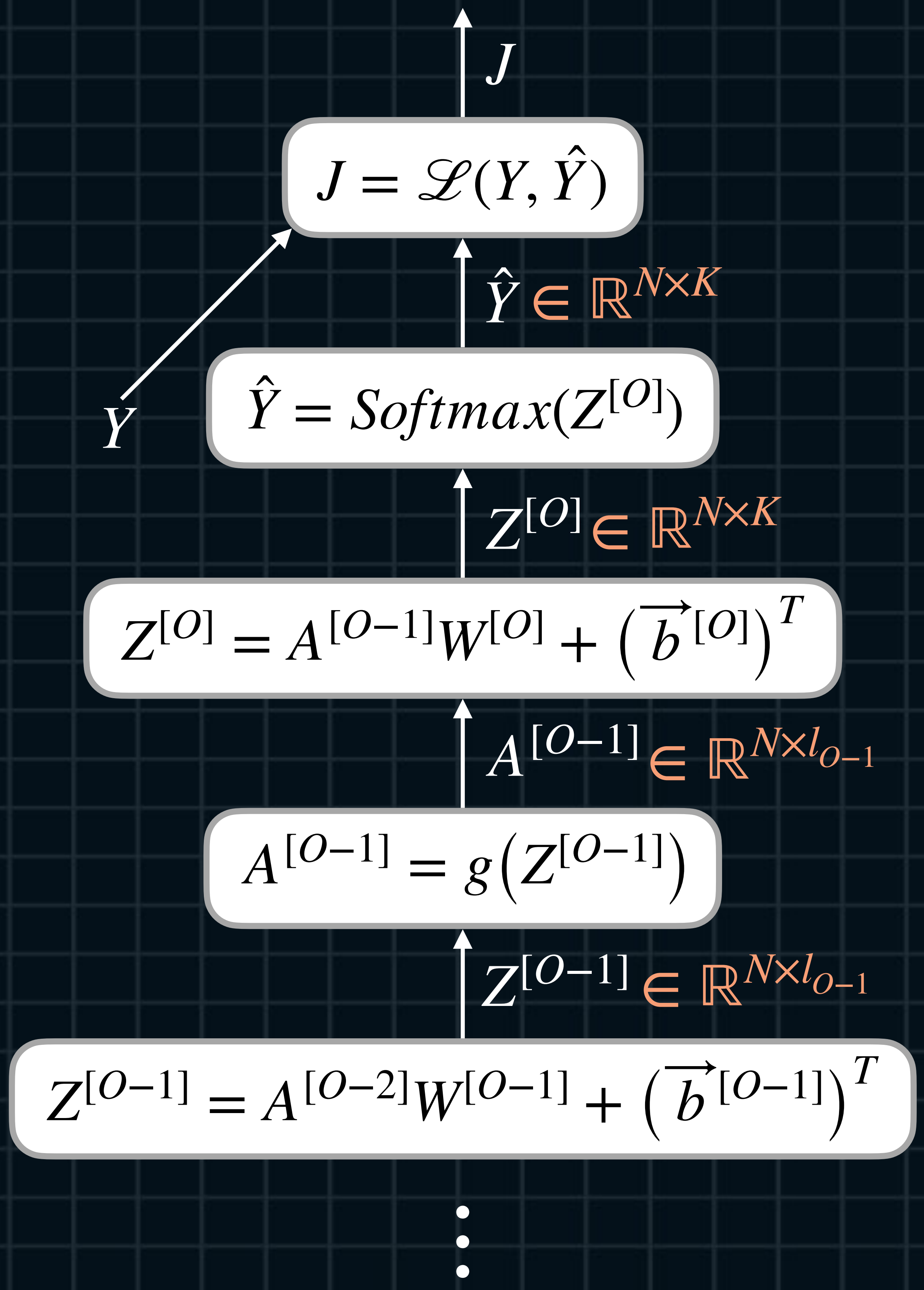


$$\hat{Y} = \begin{pmatrix} \longleftarrow & \left(\vec{\hat{y}}^{(1)}\right)^T & \longrightarrow \\ \longleftarrow & \left(\vec{\hat{y}}^{(2)}\right)^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \left(\vec{\hat{y}}^{(N)}\right)^T & \longrightarrow \end{pmatrix} = \begin{pmatrix} \hat{y}_1^{(1)} & \hat{y}_2^{(1)} & \dots & \hat{y}_K^{(1)} \\ \hat{y}_1^{(2)} & \hat{y}_2^{(2)} & \dots & \hat{y}_K^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_1^{(N)} & \hat{y}_2^{(N)} & \dots & \hat{y}_K^{(N)} \end{pmatrix}$$

$$d\hat{Y} = \begin{pmatrix} \longleftarrow & \left(d\vec{\hat{y}}^{(1)}\right)^T & \longrightarrow \\ \longleftarrow & \left(d\vec{\hat{y}}^{(2)}\right)^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \left(d\vec{\hat{y}}^{(N)}\right)^T & \longrightarrow \end{pmatrix} = \begin{pmatrix} d\hat{y}_1^{(1)} & d\hat{y}_2^{(1)} & \dots & d\hat{y}_K^{(1)} \\ d\hat{y}_1^{(2)} & d\hat{y}_2^{(2)} & \dots & d\hat{y}_K^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ d\hat{y}_1^{(N)} & d\hat{y}_2^{(N)} & \dots & d\hat{y}_K^{(N)} \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial \hat{y}_1^{(1)}} & \frac{\partial J}{\partial \hat{y}_2^{(1)}} & \dots & \frac{\partial J}{\partial \hat{y}_K^{(1)}} \\ \frac{\partial J}{\partial \hat{y}_1^{(2)}} & \frac{\partial J}{\partial \hat{y}_2^{(2)}} & \dots & \frac{\partial J}{\partial \hat{y}_K^{(2)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial \hat{y}_1^{(N)}} & \frac{\partial J}{\partial \hat{y}_2^{(N)}} & \dots & \frac{\partial J}{\partial \hat{y}_K^{(N)}} \end{pmatrix}$$

$$d\hat{Y} \in \mathbb{R}^{N \times K}$$

Jacobians within Neural Networks



$$Z^{[O]} = \begin{pmatrix} \longleftarrow & \left(\vec{z}^{[O](1)}\right)^T & \longrightarrow \\ \longleftarrow & \left(\vec{z}^{[O](2)}\right)^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \left(\vec{z}^{[O](N)}\right)^T & \longrightarrow \end{pmatrix} = \begin{pmatrix} z_1^{[O](1)} & z_2^{[O](1)} & \cdots & z_K^{[O](1)} \\ z_1^{[O](2)} & z_2^{[O](2)} & \cdots & z_K^{[O](2)} \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{[O](N)} & z_2^{[O](N)} & \cdots & z_K^{[O](N)} \end{pmatrix}$$

$$dZ^{[O]} = \begin{pmatrix} \longleftarrow & \left(d\vec{z}^{[O](1)}\right)^T & \longrightarrow \\ \longleftarrow & \left(d\vec{z}^{[O](2)}\right)^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \left(d\vec{z}^{[O](N)}\right)^T & \longrightarrow \end{pmatrix} = \begin{pmatrix} dz_1^{[O](1)} & dz_2^{[O](1)} & \cdots & dz_K^{[O](1)} \\ dz_1^{[O](2)} & dz_2^{[O](2)} & \cdots & dz_K^{[O](2)} \\ \vdots & \vdots & \ddots & \vdots \\ dz_1^{[O](N)} & dz_2^{[O](N)} & \cdots & dz_K^{[O](N)} \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial z_1^{[O](1)}} & \frac{\partial J}{\partial z_2^{[O](1)}} & \cdots & \frac{\partial J}{\partial z_K^{[O](1)}} \\ \frac{\partial J}{\partial z_1^{[O](2)}} & \frac{\partial J}{\partial z_2^{[O](2)}} & \cdots & \frac{\partial J}{\partial z_K^{[O](2)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial z_1^{[O](N)}} & \frac{\partial J}{\partial z_2^{[O](N)}} & \cdots & \frac{\partial J}{\partial z_K^{[O](N)}} \end{pmatrix}$$

$dZ^{[O]} \in \mathbb{R}^{N \times K}$

Lecture.9 Expansion of Jacobians - New Notations

Jacobians within Neural Networks

$$J$$

$$J = \mathscr{L}(Y, \hat{Y})$$

$$Y$$

$$\hat{Y} \in \mathbb{R}^{N\times K}$$

$$\hat{Y} = Softmax(Z^{[O]})$$

$$Z^{[O]} \in \mathbb{R}^{N\times K}$$

$$Z^{[O]} = A^{[O-1]}W^{[O]} + \left(\vec{b}^{[O]}\right)^T$$

$$A^{[O-1]} \in \mathbb{R}^{N\times l_{O-1}}$$

$$A^{[O-1]} = g\left(Z^{[O-1]}\right)$$

$$Z^{[O-1]} \in \mathbb{R}^{N\times l_{O-1}}$$

$$Z^{[O-1]} = A^{[O-2]}W^{[O-1]} + \left(\vec{b}^{[O-1]}\right)^T$$

$$\vdots$$

$$W^{[O]} = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \vec{w}_1 & \vec{w}_2 & \cdots & \vec{w}_K \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,K} \\ w_{2,1} & w_{2,2} & \dots & w_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l_{O-1},1} & w_{l_{O-1},2} & \dots & w_{l_{O-1},K} \end{pmatrix} \in \mathbb{R}^{l_{O-1}\times l_O} = \mathbb{R}^{l_{O-1}\times K}$$

$$dW^{[O]} = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ d\vec{w}_1^{[O]} & d\vec{w}_2^{[O]} & \cdots & d\vec{w}_K^{[O]} \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} dw_{1,1}^{[O]} & dw_{1,2}^{[O]} & \dots & dw_{1,K}^{[O]} \\ dw_{2,1}^{[O]} & dw_{2,2}^{[O]} & \dots & dw_{2,K}^{[O]} \\ \vdots & \vdots & \ddots & \vdots \\ dw_{l_{O-1},1}^{[O]} & dw_{l_{O-1},2}^{[O]} & \dots & dw_{l_{O-1},K}^{[O]} \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial w_{1,1}^{[O]}} & \frac{\partial J}{\partial w_{1,2}^{[O]}} & \cdots & \frac{\partial J}{\partial w_{1,K}^{[O]}} \\ \frac{\partial J}{\partial w_{2,1}^{[O]}} & \frac{\partial J}{\partial w_{2,2}^{[O]}} & \cdots & \frac{\partial J}{\partial w_{2,K}^{[O]}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial J}{\partial w_{l_{O-1},1}^{[O]}} & \frac{\partial J}{\partial w_{l_{O-1},2}^{[O]}} & \cdots & \frac{\partial J}{\partial w_{l_{O-1},K}^{[O]}} \end{pmatrix}$$

$$dW^{[O]} \in \mathbb{R}^{l_{O-1}\times K}$$

$$\vec{b}^{[O]} = \begin{pmatrix} b_1^{[O]} \\ b_2^{[O]} \\ \vdots \\ b_K^{[O]} \end{pmatrix} \in \mathbb{R}^K \qquad d\vec{b}^{[O]} = \begin{pmatrix} db_1^{[O]} \\ db_2^{[O]} \\ \vdots \\ db_K^{[O]} \end{pmatrix} = \begin{pmatrix} \frac{\partial J}{\partial b_1^{[O]}} \\ \frac{\partial J}{\partial b_2^{[O]}} \\ \vdots \\ \frac{\partial J}{\partial b_K^{[O]}} \end{pmatrix} \in \mathbb{R}^K$$

Lecture.9 Expansion of Jacobians - New Notations

Parameter Update with Expanded Jaocians

J

$J = \mathscr{L}(Y, \hat{Y})$

$\hat{Y} \in \mathbb{R}^{N \times K}$

Y

$\hat{Y} = Softmax(Z^{[O]})$

$Z^{[O]} \in \mathbb{R}^{N \times K}$

$Z^{[O]} = A^{[O-1]} W^{[O]} + \left(\overrightarrow{b}^{[O]}\right)^T$

$A^{[O-1]} \in \mathbb{R}^{N \times l_{O-1}}$

$A^{[O-1]} = g\left(Z^{[O-1]}\right)$

$Z^{[O-1]} \in \mathbb{R}^{N \times l_{O-1}}$

$Z^{[O-1]} = A^{[O-2]} W^{[O-1]} + \left(\overrightarrow{b}^{[O-1]}\right)^T$

$\vdots$

$$W^{[O]} := W^{[O]} - \alpha \cdot dW^{[O]}$$
$$\overrightarrow{b}^{[O]} := \overrightarrow{b}^{[O]} - \alpha \cdot d\overrightarrow{b}^{[O]}$$

$$W^{[O-1]} := W^{[O-1]} - \alpha \cdot dW^{[O-1]}$$
$$\overrightarrow{b}^{[O-1]} := \overrightarrow{b}^{[O-1]} - \alpha \cdot d\overrightarrow{b}^{[O-1]}$$

$$\vdots$$

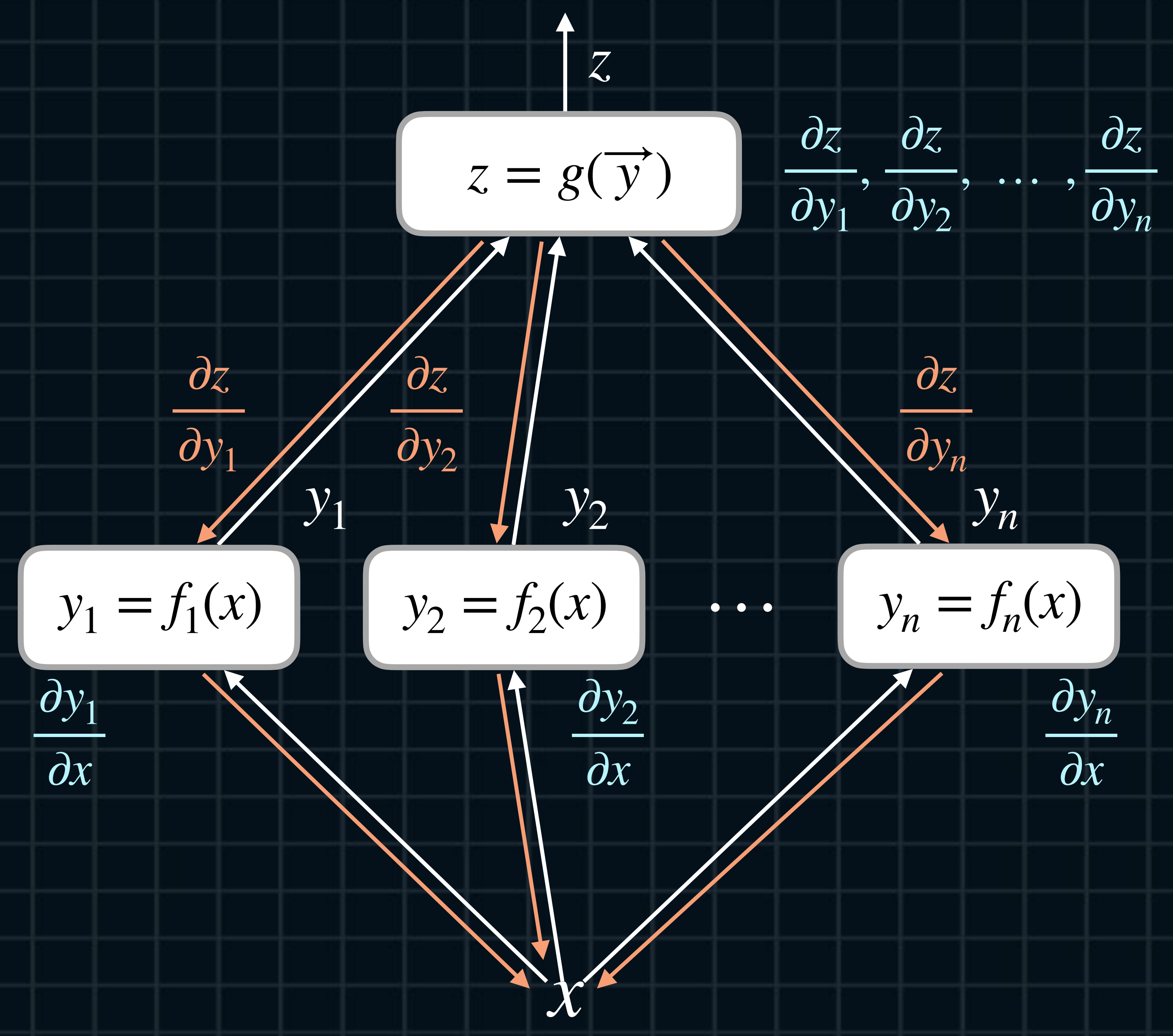
$$W^{[1]} := W^{[1]} - \alpha \cdot dW^{[1]}$$
$$\overrightarrow{b}^{[1]} := \overrightarrow{b}^{[1]} - \alpha \cdot d\overrightarrow{b}^{[1]}$$

- New Notations

Matrix-Matrix Jacobians

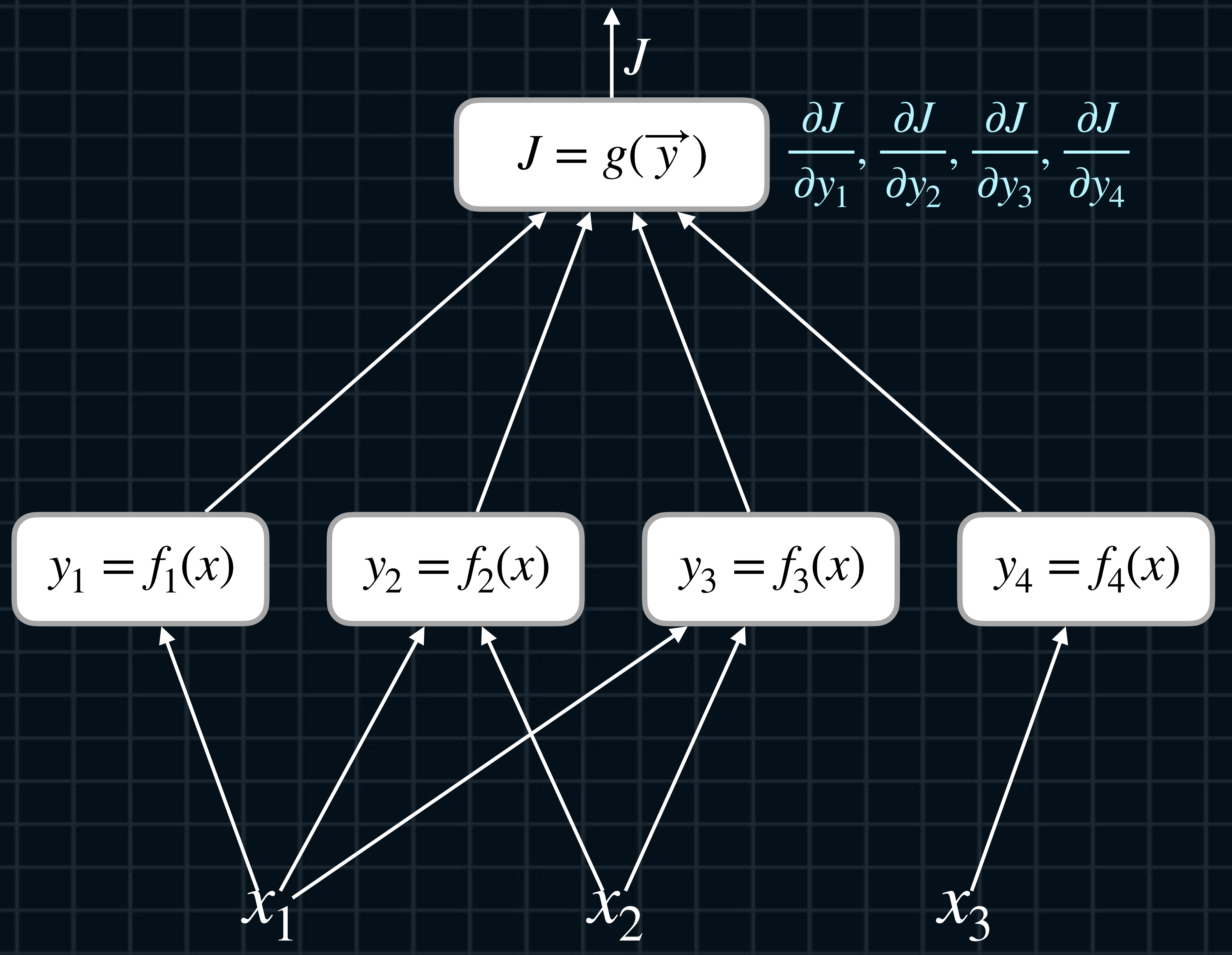
$$\frac{\partial J}{\partial Y}\frac{\partial Y}{\partial X}$$

Effect & Total Derivatives



Lecture.9 Expansion of Jacobians

- Keypoints

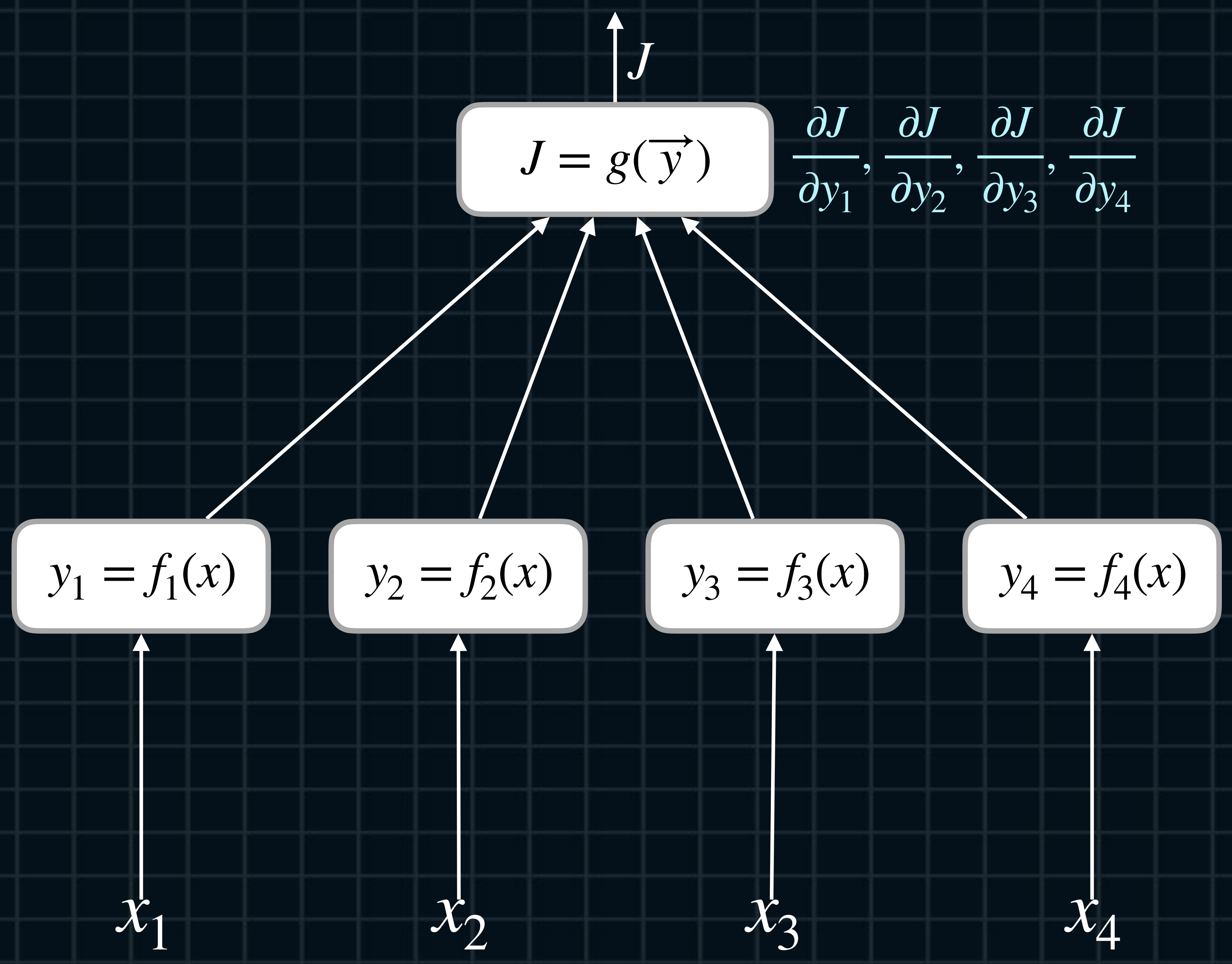


$$\frac{\partial J}{\partial x_1} = \frac{\partial J}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial J}{\partial y_2}\frac{\partial y_2}{\partial x_1} + \frac{\partial J}{\partial y_3}\frac{\partial y_3}{\partial x_1}$$

$$\frac{\partial J}{\partial x_2} = \frac{\partial J}{\partial y_2}\frac{\partial y_2}{\partial x_2} + \frac{\partial J}{\partial y_3}\frac{\partial y_3}{\partial x_2}$$

$$\frac{\partial J}{\partial x_3} = \frac{\partial J}{\partial y_4}\frac{\partial y_4}{\partial x_3}$$

Lecture.9
Expansion of Jacobians - Keypoints



$$\frac{\partial J}{\partial x_1} = \frac{\partial J}{\partial y_1}\frac{\partial y_1}{\partial x_1}$$

$$\frac{\partial J}{\partial x_3} = \frac{\partial J}{\partial y_3}\frac{\partial y_3}{\partial x_3}$$

$$\frac{\partial J}{\partial x_2} = \frac{\partial J}{\partial y_2}\frac{\partial y_2}{\partial x_2}$$

$$\frac{\partial J}{\partial x_4} = \frac{\partial J}{\partial y_4}\frac{\partial y_4}{\partial x_4}$$

Lecture.9 Expansion of Jacobians - Simple Examples

Sum Operation

$$X \in \mathbb{R}^{\alpha \times \beta},\ J \in \mathbb{R}$$

$$J = \sum_{r=1}^{\alpha} \sum_{c=1}^{\beta} x_{rc}$$

$$\frac{\partial J}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \left[\sum_{r=1}^{\alpha} \sum_{c=1}^{\beta} x_{rc} \right] = 1$$

$$\frac{\partial J}{\partial X} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{12}} & \dots & \frac{\partial y}{\partial x_{1n}} \\ \frac{\partial y}{\partial x_{21}} & \frac{\partial y}{\partial x_{22}} & \dots & \frac{\partial y}{\partial x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y}{\partial x_{m1}} & \frac{\partial y}{\partial x_{m2}} & \dots & \frac{\partial y}{\partial x_{mn}} \end{pmatrix} = \begin{pmatrix} dx_{11} & dx_{12} & \dots & dx_{1n} \\ dx_{21} & dx_{22} & \dots & dx_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ dx_{m1} & dx_{m2} & \dots & dx_{mn} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Lecture.9 Expansion of Jacobians - Simple Examples

Mean Operation

$$X \in \mathbb{R}^{\alpha \times \beta},\ J \in \mathbb{R}$$

$$J = \frac{1}{\alpha\beta} \sum_{r=1}^{\alpha} \sum_{c=1}^{\beta} x_{rc}$$

$$\frac{\partial J}{\partial x_{ij}} = \frac{\partial}{\partial x_{ij}} \left[\frac{1}{\alpha\beta} \sum_{r=1}^{\alpha} \sum_{c=1}^{\beta} x_{rc} \right] = \frac{1}{\alpha\beta}$$

$$\frac{\partial y}{\partial X} = \begin{pmatrix} \frac{1}{\alpha\beta} & \frac{1}{\alpha\beta} & \cdots & \frac{1}{\alpha\beta} \\ \frac{1}{\alpha\beta} & \frac{1}{\alpha\beta} & \cdots & \frac{1}{\alpha\beta} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\alpha\beta} & \frac{1}{\alpha\beta} & \cdots & \frac{1}{\alpha\beta} \end{pmatrix} = \frac{1}{\alpha\beta} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

Lecture.9
Expansion of Jacobians

- Unary Element-wise Operations

General Case

$$X, Y \in \mathbb{R}^{\alpha\times\beta}$$

$$Y = f(X), y_{ij} = f(x_{ij})$$

$$dY = \begin{pmatrix} dy_{11} & dy_{12} & \dots & dy_{1\beta} \\ dy_{21} & dy_{22} & \dots & dy_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} & dy_{\alpha 2} & \dots & dy_{\alpha\beta} \end{pmatrix} \longrightarrow dX = \begin{pmatrix} dx_{11} & dx_{12} & \dots & dx_{1\beta} \\ dx_{21} & dx_{22} & \dots & dx_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dx_{\alpha 1} & dx_{\alpha 2} & \dots & dx_{\alpha\beta} \end{pmatrix}$$

- Unary Element-wise Operations

General Case

$$X, Y \in \mathbb{R}^{\alpha\times\beta}$$
$$Y = f(X), y_{ij} = f(x_{ij})$$

$y_{12} = f(x_{12})$

$$\begin{pmatrix} x_{11} & \boxed{x_{12}} & \cdots & x_{1\beta} \\ x_{21} & x_{22} & \cdots & \boxed{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{x_{\alpha 1}} & x_{\alpha 2} & \cdots & x_{\alpha\beta} \end{pmatrix} \quad \begin{pmatrix} y_{11} & \boxed{y_{12}} & \cdots & y_{1\beta} \\ y_{21} & y_{22} & \cdots & \boxed{y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{y_{\alpha 1}} & y_{\alpha 2} & \cdots & y_{\alpha\beta} \end{pmatrix}$$

$y_{2\beta} = f(x_{2\beta})$

$y_{\alpha 1} = f(x_{\alpha 1})$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \cdots & \frac{\partial y_{1\beta}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix}$$

Lecture.9 Expansion of Jacobians - Unary Element-wise Operations

General Case

$$X, Y \in \mathbb{R}^{\alpha\times\beta}$$

$$Y = f(X), y_{ij} = f(x_{ij})$$

$$\frac{\partial J}{\partial y_{2\beta}}\frac{\partial y_{2\beta}}{\partial x_{2\beta}}$$

$$\frac{\partial J}{\partial y_{12}}\frac{\partial y_{12}}{\partial x_{12}} \qquad y_{12} = f(x_{12})$$

$$\begin{pmatrix} x_{11} & \boxed{x_{12}} & \cdots & x_{1\beta} \\ x_{21} & x_{22} & \cdots & \boxed{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{x_{\alpha 1}} & x_{\alpha 2} & \cdots & x_{\alpha\beta} \end{pmatrix} \quad \begin{pmatrix} y_{11} & \boxed{y_{12}} & \cdots & y_{1\beta} \\ y_{21} & y_{22} & \cdots & \boxed{y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{y_{\alpha 1}} & y_{\alpha 2} & \cdots & y_{\alpha\beta} \end{pmatrix} \quad y_{2\beta} = f(x_{2\beta})$$

$$\frac{\partial J}{\partial y_{\alpha 1}}\frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} \qquad y_{\alpha 1} = f(x_{\alpha 1})$$

$$dx_{ij} = dy_{ij}\frac{\partial y_{ij}}{\partial x_{ij}}$$

- Unary Element-wise Operations

General Case

$$dx_{ij} = dy_{ij}\frac{\partial y_{ij}}{\partial x_{ij}}$$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \cdots & \frac{\partial y_{1\beta}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix}$$

$$dX = dY\frac{\partial Y}{\partial X} = \begin{pmatrix} dy_{11}\cdot\frac{\partial y_{11}}{\partial x_{11}} & dy_{12}\cdot\frac{\partial y_{12}}{\partial x_{12}} & \cdots & dy_{1\beta}\cdot\frac{\partial y_{1\beta}}{\partial x_{1\beta}} \\ dy_{21}\cdot\frac{\partial y_{21}}{\partial x_{21}} & dy_{22}\cdot\frac{\partial y_{22}}{\partial x_{22}} & \cdots & dy_{2\beta}\cdot\frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1}\cdot\frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & dy_{\alpha 2}\cdot\frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & dy_{\alpha\beta}\cdot\frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = dY \bigcirc \frac{\partial Y}{\partial X}$$

Lecture.9 Expansion of Jacobians - Unary Element-wise Operations

General Case

$$\frac{\partial J}{\partial y_{12}} \frac{\partial y_{12}}{\partial x_{12}} \qquad \frac{\partial J}{\partial y_{2\beta}} \frac{\partial y_{2\beta}}{\partial x_{2\beta}}$$

$$y_{12} = f(x_{12})$$

$$\begin{pmatrix} x_{11} & \boxed{x_{12}} & \cdots & x_{1\beta} \\ x_{21} & x_{22} & \cdots & \boxed{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{x_{\alpha 1}} & x_{\alpha 2} & \cdots & x_{\alpha\beta} \end{pmatrix} \qquad \begin{pmatrix} y_{11} & \boxed{y_{12}} & \cdots & y_{1\beta} \\ y_{21} & y_{22} & \cdots & \boxed{y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{y_{\alpha 1}} & y_{\alpha 2} & \cdots & y_{\alpha\beta} \end{pmatrix} \quad y_{2\beta} = f(x_{2\beta})$$

$$\frac{\partial J}{\partial y_{\alpha 1}} \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} \qquad y_{\alpha 1} = f(x_{\alpha 1})$$

$$dX = dY \circ \frac{\partial Y}{\partial X}$$

- Unary Element-wise Operations

Basic Examples

$$Y = e^X, y_{ij} = e^{ij}$$

$$\frac{\partial y_{ij}}{\partial x_{ij}} = e^{x_{ij}} = y_{ij}$$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \cdots & \frac{\partial y_{12}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1\beta} \\ y_{21} & y_{22} & \cdots & y_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ y_{\alpha 1} & y_{\alpha 2} & \cdots & y_{\alpha\beta} \end{pmatrix} = Y$$

$$dJ = dY \frac{\partial Y}{\partial X} = \begin{pmatrix} dy_{11} \cdot y_{11} & dy_{12} \cdot y_{12} & \cdots & dy_{1\beta} \cdot y_{1\beta} \\ dy_{21} \cdot y_{21} & dy_{22} \cdot y_{22} & \cdots & dy_{2\beta} \cdot y_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot y_{\alpha 1} & dy_{\alpha 2} \cdot y_{\alpha 2} & \cdots & dy_{\alpha\beta} \cdot y_{\alpha\beta} \end{pmatrix} = \frac{\partial J}{\partial Y} \bigcirc Y$$

Lecture.9 Expansion of Jacobians - Unary Element-wise Operations

Basic Examples

$$Y = log(X),\ y_{ij} = log(x_{ij})$$

$$\frac{\partial y_{ij}}{\partial x_{ij}} = \frac{1}{x_{ij}}$$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \cdots & \frac{\partial y_{12}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} \frac{1}{x_{11}} & \frac{1}{x_{12}} & \cdots & \frac{1}{x_{1\beta}} \\ \frac{1}{x_{21}} & \frac{1}{x_{22}} & \cdots & \frac{1}{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_{\alpha 1}} & \frac{1}{x_{\alpha 2}} & \cdots & \frac{1}{x_{\alpha\beta}} \end{pmatrix} = \frac{1}{X}$$

$$dX = dY \frac{\partial Y}{\partial X} = \begin{pmatrix} dy_{11} \cdot \frac{1}{x_{11}} & dy_{12} \cdot \frac{1}{x_{12}} & \cdots & dy_{1\beta} \cdot \frac{1}{x_{1\beta}} \\ dy_{21} \cdot \frac{1}{x_{21}} & dy_{22} \cdot \frac{1}{x_{22}} & \cdots & dy_{2\beta} \cdot \frac{1}{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \frac{1}{x_{\alpha 1}} & dy_{\alpha 2} \cdot \frac{1}{x_{\alpha 2}} & \cdots & dy_{\alpha\beta} \cdot \frac{1}{x_{\alpha\beta}} \end{pmatrix} = \frac{\partial J}{\partial Y} \bigcirc \frac{1}{X}$$

Lecture.9 Expansion of Jacobians - Unary Element-wise Operations

Activation Functions

$$Y = \sigma(X)$$

$$\frac{\partial y_{ij}}{\partial x_{ij}} = y_{ij}(1 - y_{ij})$$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \dots & \frac{\partial y_{12}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \dots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \dots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} y_{11}(1 - y_{11}) & y_{12}(1 - y_{12}) & \dots & y_{1\beta}(1 - y_{1\beta}) \\ y_{21}(1 - y_{21}) & y_{22}(1 - y_{22}) & \dots & y_{2\beta}(1 - y_{2\beta}) \\ \vdots & \vdots & \ddots & \vdots \\ y_{\alpha 1}(1 - y_{\alpha 1}) & y_{\alpha 2}(1 - y_{\alpha 2}) & \dots & y_{\alpha\beta}(1 - y_{\alpha\beta}) \end{pmatrix} = Y \bigcirc \left(1 - Y\right)$$

$$dX = dY\frac{\partial Y}{\partial X} = \begin{pmatrix} dy_{11} \cdot y_{11}(1 - y_{11}) & dy_{12} \cdot y_{12}(1 - y_{12}) & \dots & dy_{1\beta} \cdot y_{1\beta}(1 - y_{1\beta}) \\ dy_{21} \cdot y_{21}(1 - y_{21}) & dy_{22} \cdot y_{22}(1 - y_{22}) & \dots & dy_{2\beta} \cdot y_{2\beta}(1 - y_{2\beta}) \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot y_{\alpha 1}(1 - y_{\alpha 1}) & dy_{\alpha 2} \cdot y_{\alpha 2}(1 - y_{\alpha 2}) & \dots & dy_{\alpha\beta} \cdot y_{\alpha\beta}(1 - y_{\alpha\beta}) \end{pmatrix} = dY \bigcirc \left(Y \bigcirc (1 - Y)\right)$$

Activation Functions

$$Y = tanh(X)$$

$$\frac{\partial y_{ij}}{\partial x_{ij}} = (1 + y_{ij})(1 - y_{ij})$$

$$\frac{\partial Y}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial x_{11}} & \frac{\partial y_{12}}{\partial x_{12}} & \dots & \frac{\partial y_{12}}{\partial x_{1\beta}} \\ \frac{\partial y_{21}}{\partial x_{21}} & \frac{\partial y_{22}}{\partial x_{22}} & \dots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial y_{\alpha 2}}{\partial x_{\alpha 2}} & \dots & \frac{\partial y_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} (1 + y_{11})(1 - y_{11}) & (1 + y_{12})(1 - y_{12}) & \dots & (1 + y_{1\beta})(1 - y_{1\beta}) \\ (1 + y_{21})(1 - y_{21}) & (1 + y_{22})(1 - y_{22}) & \dots & (1 + y_{2\beta})(1 - y_{2\beta}) \\ \vdots & \vdots & \ddots & \vdots \\ (1 + y_{\alpha 1})(1 - y_{\alpha 1}) & (1 + y_{\alpha 2})(1 - y_{\alpha 2}) & \dots & (1 + y_{\alpha\beta})(1 - y_{\alpha\beta}) \end{pmatrix} = \left(1 + Y\right) \bigcirc \left(1 - Y\right)$$

$$dX = dY \frac{\partial Y}{\partial X} = \begin{pmatrix} dy_{11} \cdot (1 + y_{11})(1 - y_{11}) & dy_{12} \cdot (1 + y_{12})(1 - y_{12}) & \dots & dy_{1\beta} \cdot (1 + y_{1\beta})(1 - y_{1\beta}) \\ dy_{21} \cdot (1 + y_{21})(1 - y_{21}) & dy_{22} \cdot (1 + y_{22})(1 - y_{22}) & \dots & dy_{2\beta} \cdot (1 + y_{2\beta})(1 - y_{2\beta}) \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot (1 + y_{\alpha 1})(1 - y_{\alpha 1}) & dy_{\alpha 2} \cdot (1 + y_{\alpha 2})(1 - y_{\alpha 2}) & \dots & dy_{\alpha\beta} \cdot (1 + y_{\alpha\beta})(1 - y_{\alpha\beta}) \end{pmatrix} = dY \bigcirc \Big((1 + Y) \bigcirc (1 - Y) \Big)$$

Lecture.9 Expansion of Jacobians - Unary Element-wise Operations

Activation Functions

$$Y = ReLU(X)$$

$$\frac{\partial y_{ij}}{\partial x_{ij}} = \begin{cases} 1, & x_{ij} \geq 0 \\ 0, & x_{ij} < 0 \end{cases}$$

$$dX = dY\frac{\partial Y}{\partial X} = = \begin{pmatrix} dy_{11} \cdot \begin{cases} 1, & x_{11} \geq 0 \\ 0, & x_{11} < 0 \end{cases} & dy_{12} \cdot \begin{cases} 1, & x_{12} \geq 0 \\ 0, & x_{12} < 0 \end{cases} & \dots & dy_{1\beta} \cdot \begin{cases} 1, & x_{1\beta} \geq 0 \\ 0, & x_{1\beta} < 0 \end{cases} \\ dy_{21} \cdot \begin{cases} 1, & x_{21} \geq 0 \\ 0, & x_{21} < 0 \end{cases} & dy_{22} \cdot \begin{cases} 1, & x_{22} \geq 0 \\ 0, & x_{22} < 0 \end{cases} & \dots & dy_{2\beta} \cdot \begin{cases} 1, & x_{2\beta} \geq 0 \\ 0, & x_{2\beta} < 0 \end{cases} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \begin{cases} 1, & x_{\alpha 1} \geq 0 \\ 0, & x_{\alpha 1} < 0 \end{cases} & dy_{\alpha 2} \cdot \begin{cases} 1, & x_{\alpha 2} \geq 0 \\ 0, & x_{\alpha 2} < 0 \end{cases} & \dots & dy_{\alpha\beta} \cdot \begin{cases} 1, & x_{\alpha\beta} \geq 0 \\ 0, & x_{\alpha\beta} < 0 \end{cases} \end{pmatrix} = dY \bigcirc (Z \geq 0).astype(float)$$

Lecture.9 Expansion of Jacobians - Binary Element-wise Operations

General Case

$$X, Y, Z \in \mathbb{R}^{\alpha \times \beta}$$

$$Z = f(X, Y),\ z_{ij} = f(x_{ij}, y_{ij})$$

$$\begin{pmatrix} x_{11} & \boxed{x_{12}} & \cdots & x_{1\beta} \\ x_{21} & x_{22} & \cdots & \boxed{x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{x_{\alpha 1}} & x_{\alpha 2} & \cdots & x_{\alpha\beta} \end{pmatrix} \qquad \begin{pmatrix} z_{11} & \boxed{z_{12}} & \cdots & z_{1\beta} \\ z_{21} & z_{22} & \cdots & \boxed{z_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{z_{\alpha 1}} & z_{\alpha 2} & \cdots & z_{\alpha\beta} \end{pmatrix}$$

$$\begin{pmatrix} y_{11} & \boxed{y_{12}} & \cdots & y_{1\beta} \\ y_{21} & y_{22} & \cdots & \boxed{y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \boxed{y_{\alpha 1}} & y_{\alpha 2} & \cdots & y_{\alpha\beta} \end{pmatrix}$$

$$\frac{\partial J}{\partial x_{ij}} = \frac{\partial J}{\partial z_{ij}} \frac{\partial z_{ij}}{\partial x_{ij}} \qquad \frac{\partial J}{\partial y_{ij}} = \frac{\partial J}{\partial z_{ij}} \frac{\partial z_{ij}}{\partial y_{ij}}$$

$$dx_{ij} = dz_{ij} \frac{\partial z_{ij}}{\partial x_{ij}} \qquad dy_{ij} = dz_{ij} \frac{\partial z_{ij}}{\partial y_{ij}}$$

- Binary Element-wise Operations

General Case

$$Z = f(X, Y),\ z_{ij} = f(x_{ij}, y_{ij})$$

$$dx_{ij} = dz_{ij}\frac{\partial z_{ij}}{\partial x_{ij}} \qquad dy_{ij} = dz_{ij}\frac{\partial z_{ij}}{\partial y_{ij}}$$

$$\frac{\partial Z}{\partial X} = \begin{pmatrix} \frac{\partial z_{11}}{\partial x_{11}} & \frac{\partial z_{12}}{\partial x_{12}} & \cdots & \frac{\partial z_{12}}{\partial x_{1\beta}} \\ \frac{\partial z_{21}}{\partial x_{21}} & \frac{\partial z_{22}}{\partial x_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_{\alpha 1}}{\partial x_{\alpha 1}} & \frac{\partial z_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & \frac{\partial z_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} \qquad \frac{\partial Z}{\partial Y} = \begin{pmatrix} \frac{\partial z_{11}}{\partial y_{11}} & \frac{\partial z_{12}}{\partial y_{12}} & \cdots & \frac{\partial z_{12}}{\partial y_{1\beta}} \\ \frac{\partial z_{21}}{\partial y_{21}} & \frac{\partial z_{22}}{\partial y_{22}} & \cdots & \frac{\partial y_{2\beta}}{\partial y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_{\alpha 1}}{\partial y_{\alpha 1}} & \frac{\partial z_{\alpha 2}}{\partial y_{\alpha 2}} & \cdots & \frac{\partial z_{\alpha\beta}}{\partial y_{\alpha\beta}} \end{pmatrix}$$

$$dX = dZ\frac{\partial Z}{\partial X} = \begin{pmatrix} dz_{11}\cdot\frac{\partial z_{11}}{\partial x_{11}} & dz_{12}\cdot\frac{\partial z_{12}}{\partial x_{12}} & \cdots & dz_{1\beta}\cdot\frac{\partial z_{1\beta}}{\partial x_{1\beta}} \\ dz_{21}\cdot\frac{\partial z_{21}}{\partial x_{21}} & dz_{22}\cdot\frac{\partial z_{22}}{\partial x_{22}} & \cdots & dz_{2\beta}\cdot\frac{\partial z_{2\beta}}{\partial x_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dz_{\alpha 1}\cdot\frac{\partial z_{\alpha 1}}{\partial x_{\alpha 1}} & dz_{\alpha 2}\cdot\frac{\partial z_{\alpha 2}}{\partial x_{\alpha 2}} & \cdots & dz_{\alpha\beta}\cdot\frac{\partial z_{\alpha\beta}}{\partial x_{\alpha\beta}} \end{pmatrix} = dZ \bigcirc \frac{\partial Z}{\partial X}$$

$$dY = dZ\frac{\partial Z}{\partial Y} = \begin{pmatrix} dz_{11}\cdot\frac{\partial z_{11}}{\partial y_{11}} & dz_{12}\cdot\frac{\partial z_{12}}{\partial y_{12}} & \cdots & dz_{1\beta}\cdot\frac{\partial z_{1\beta}}{\partial y_{1\beta}} \\ dz_{21}\cdot\frac{\partial z_{21}}{\partial y_{21}} & dz_{22}\cdot\frac{\partial z_{22}}{\partial y_{22}} & \cdots & dz_{2\beta}\cdot\frac{\partial z_{2\beta}}{\partial y_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dz_{\alpha 1}\cdot\frac{\partial z_{\alpha 1}}{\partial y_{\alpha 1}} & dz_{\alpha 2}\cdot\frac{\partial z_{\alpha 2}}{\partial y_{\alpha 2}} & \cdots & dz_{\alpha\beta}\cdot\frac{\partial z_{\alpha\beta}}{\partial y_{\alpha\beta}} \end{pmatrix} = dZ \bigcirc \frac{\partial Z}{\partial Y}$$

Lecture.9 Expansion of Jacobians - Binary Element-wise Operations

Basic Examples

$$Y = M + N$$

$$y_{ij} = m_{ij} + n_{ij} \qquad \frac{\partial y_{ij}}{\partial m_{ij}} = \frac{\partial y_{ij}}{\partial n_{ij}} = 1$$

$$dM = \begin{pmatrix} dy_{11} \cdot \frac{\partial y_{11}}{\partial m_{11}} & dy_{12} \cdot \frac{\partial y_{12}}{\partial m_{12}} & \dots & dy_{1\beta} \cdot \frac{\partial y_{12}}{\partial m_{1\beta}} \\ dy_{21} \cdot \frac{\partial y_{21}}{\partial m_{21}} & dy_{22} \cdot \frac{\partial y_{22}}{\partial m_{22}} & \dots & dy_{2\beta} \cdot \frac{\partial y_{2\beta}}{\partial m_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \frac{\partial y_{\alpha 1}}{\partial m_{\alpha 1}} & dy_{\alpha 2} \cdot \frac{\partial y_{\alpha 2}}{\partial m_{\alpha 2}} & \dots & dy_{\alpha\beta} \cdot \frac{\partial y_{\alpha\beta}}{\partial m_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11} \cdot 1 & dy_{12} \cdot 1 & \dots & dy_{1\beta} \cdot 1 \\ dy_{21} \cdot 1 & dy_{22} \cdot 1 & \dots & dy_{2\beta} \cdot 1 \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot 1 & dy_{\alpha 2} \cdot 1 & \dots & dy_{\alpha\beta} \cdot 1 \end{pmatrix} = \begin{pmatrix} dy_{11} & dy_{12} & \dots & dy_{1\beta} \\ dy_{21} & dy_{22} & \dots & dy_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} & dy_{\alpha 2} & \dots & dy_{\alpha\beta} \end{pmatrix} = dY$$

$$dN = \begin{pmatrix} dy_{11} \cdot \frac{\partial y_{11}}{\partial n_{11}} & dy_{12} \cdot \frac{\partial y_{12}}{\partial n_{12}} & \dots & dy_{1\beta} \cdot \frac{\partial y_{12}}{\partial n_{1\beta}} \\ dy_{21} \cdot \frac{\partial y_{21}}{\partial n_{21}} & dy_{22} \cdot \frac{\partial y_{22}}{\partial n_{22}} & \dots & dy_{2\beta} \cdot \frac{\partial y_{2\beta}}{\partial n_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \frac{\partial y_{\alpha 1}}{\partial n_{\alpha 1}} & dy_{\alpha 2} \cdot \frac{\partial y_{\alpha 2}}{\partial n_{\alpha 2}} & \dots & dy_{\alpha\beta} \cdot \frac{\partial y_{\alpha\beta}}{\partial n_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11} \cdot 1 & dy_{12} \cdot 1 & \dots & dy_{1\beta} \cdot 1 \\ dy_{21} \cdot 1 & dy_{22} \cdot 1 & \dots & dy_{2\beta} \cdot 1 \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot 1 & dy_{\alpha 2} \cdot 1 & \dots & dy_{\alpha\beta} \cdot 1 \end{pmatrix} = \begin{pmatrix} dy_{11} & dy_{12} & \dots & dy_{1\beta} \\ dy_{21} & dy_{22} & \dots & dy_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} & dy_{\alpha 2} & \dots & dy_{\alpha\beta} \end{pmatrix} = dY$$

Basic Examples

$$Y = M - N$$

$$y_{ij} = m_{ij} - n_{ij} \qquad \frac{\partial y_{ij}}{\partial m_{ij}} = 1, \ \frac{\partial y_{ij}}{\partial n_{ij}} = -1$$

$$dM = \begin{pmatrix} dy_{11}\cdot\frac{\partial y_{11}}{\partial m_{11}} & dy_{12}\cdot\frac{\partial y_{12}}{\partial m_{12}} & \dots & dy_{1\beta}\cdot\frac{\partial y_{12}}{\partial m_{1\beta}} \\ dy_{21}\cdot\frac{\partial y_{21}}{\partial m_{21}} & dy_{22}\cdot\frac{\partial y_{22}}{\partial m_{22}} & \dots & dy_{2\beta}\cdot\frac{\partial y_{2\beta}}{\partial m_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1}\cdot\frac{\partial y_{\alpha 1}}{\partial m_{\alpha 1}} & dy_{\alpha 2}\cdot\frac{\partial y_{\alpha 2}}{\partial m_{\alpha 2}} & \dots & dy_{\alpha\beta}\cdot\frac{\partial y_{\alpha\beta}}{\partial m_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11}\cdot 1 & dy_{12}\cdot 1 & \dots & dy_{1\beta}\cdot 1 \\ dy_{21}\cdot 1 & dy_{22}\cdot 1 & \dots & dy_{2\beta}\cdot 1 \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1}\cdot 1 & dy_{\alpha 2}\cdot 1 & \dots & dy_{\alpha\beta}\cdot 1 \end{pmatrix} = \begin{pmatrix} dy_{11} & dy_{12} & \dots & dy_{1\beta} \\ dy_{21} & dy_{22} & \dots & dy_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} & dy_{\alpha 2} & \dots & dy_{\alpha\beta} \end{pmatrix} = dY$$

$$dN = \begin{pmatrix} dy_{11}\cdot\frac{\partial y_{11}}{\partial n_{11}} & dy_{12}\cdot\frac{\partial y_{12}}{\partial n_{12}} & \dots & dy_{1\beta}\cdot\frac{\partial y_{12}}{\partial n_{1\beta}} \\ dy_{21}\cdot\frac{\partial y_{21}}{\partial n_{21}} & dy_{22}\cdot\frac{\partial y_{22}}{\partial n_{22}} & \dots & dy_{2\beta}\cdot\frac{\partial y_{2\beta}}{\partial n_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1}\cdot\frac{\partial y_{\alpha 1}}{\partial n_{\alpha 1}} & dy_{\alpha 2}\cdot\frac{\partial y_{\alpha 2}}{\partial n_{\alpha 2}} & \dots & dy_{\alpha\beta}\cdot\frac{\partial y_{\alpha\beta}}{\partial n_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11}\cdot(-1) & dy_{12}\cdot(-1) & \dots & dy_{1\beta}\cdot(-1) \\ dy_{21}\cdot(-1) & dy_{22}\cdot(-1) & \dots & dy_{2\beta}\cdot(-1) \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1}\cdot(-1) & dy_{\alpha 2}\cdot(-1) & \dots & dy_{\alpha\beta}\cdot(-1) \end{pmatrix} = \begin{pmatrix} -dy_{11} & -dy_{12} & \dots & -dy_{1\beta} \\ -dy_{21} & -dy_{22} & \dots & -dy_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ -dy_{\alpha 1} & -dy_{\alpha 2} & \dots & -dy_{\alpha\beta} \end{pmatrix} = -dY$$

Lecture.9 Expansion of Jacobians - Binary Element-wise Operations

Basic Examples

$$Y = M \bigcirc N$$

$$y_{ij} = m_{ij} \cdot n_{ij} \qquad \frac{\partial y_{ij}}{\partial m_{ij}} = n_{ij}, \frac{\partial y_{ij}}{\partial n_{ij}} = m_{ij}$$

$$dM = \begin{pmatrix} dy_{11} \cdot \frac{\partial y_{11}}{\partial m_{11}} & dy_{12} \cdot \frac{\partial y_{12}}{\partial m_{12}} & \dots & dy_{1\beta} \cdot \frac{\partial y_{12}}{\partial m_{1\beta}} \\ dy_{21} \cdot \frac{\partial y_{21}}{\partial m_{21}} & dy_{22} \cdot \frac{\partial y_{22}}{\partial m_{22}} & \dots & dy_{2\beta} \cdot \frac{\partial y_{2\beta}}{\partial m_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \frac{\partial y_{\alpha 1}}{\partial m_{\alpha 1}} & dy_{\alpha 2} \cdot \frac{\partial y_{\alpha 2}}{\partial m_{\alpha 2}} & \dots & dy_{\alpha\beta} \cdot \frac{\partial y_{\alpha\beta}}{\partial m_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11} \cdot n_{11} & dy_{12} \cdot n_{12} & \dots & dy_{1\beta} \cdot n_{1\beta} \\ dy_{21} \cdot n_{21} & dy_{22} \cdot n_{22} & \dots & dy_{2\beta} \cdot n_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot n_{\alpha 1} & dy_{\alpha 2} \cdot n_{\alpha 2} & \dots & dy_{\alpha\beta} \cdot n_{\alpha\beta} \end{pmatrix} = dY \bigcirc N$$

$$dN = \begin{pmatrix} dy_{11} \cdot \frac{\partial y_{11}}{\partial n_{11}} & dy_{12} \cdot \frac{\partial y_{12}}{\partial n_{12}} & \dots & dy_{1\beta} \cdot \frac{\partial y_{12}}{\partial n_{1\beta}} \\ dy_{21} \cdot \frac{\partial y_{21}}{\partial n_{21}} & dy_{22} \cdot \frac{\partial y_{22}}{\partial n_{22}} & \dots & dy_{2\beta} \cdot \frac{\partial y_{2\beta}}{\partial n_{2\beta}} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot \frac{\partial y_{\alpha 1}}{\partial n_{\alpha 1}} & dy_{\alpha 2} \cdot \frac{\partial y_{\alpha 2}}{\partial n_{\alpha 2}} & \dots & dy_{\alpha\beta} \cdot \frac{\partial y_{\alpha\beta}}{\partial n_{\alpha\beta}} \end{pmatrix} = \begin{pmatrix} dy_{11} \cdot m_{11} & dy_{12} \cdot m_{12} & \dots & dy_{1\beta} \cdot m_{1\beta} \\ dy_{21} \cdot m_{21} & dy_{22} \cdot m_{22} & \dots & dy_{2\beta} \cdot m_{2\beta} \\ \vdots & \vdots & \ddots & \vdots \\ dy_{\alpha 1} \cdot m_{\alpha 1} & dy_{\alpha 2} \cdot m_{\alpha 2} & \dots & dy_{\alpha\beta} \cdot m_{\alpha\beta} \end{pmatrix} = dY \bigcirc M$$

Backpropagation and Jacobian Matrices

Lecture.9
Expansion of Jacobians