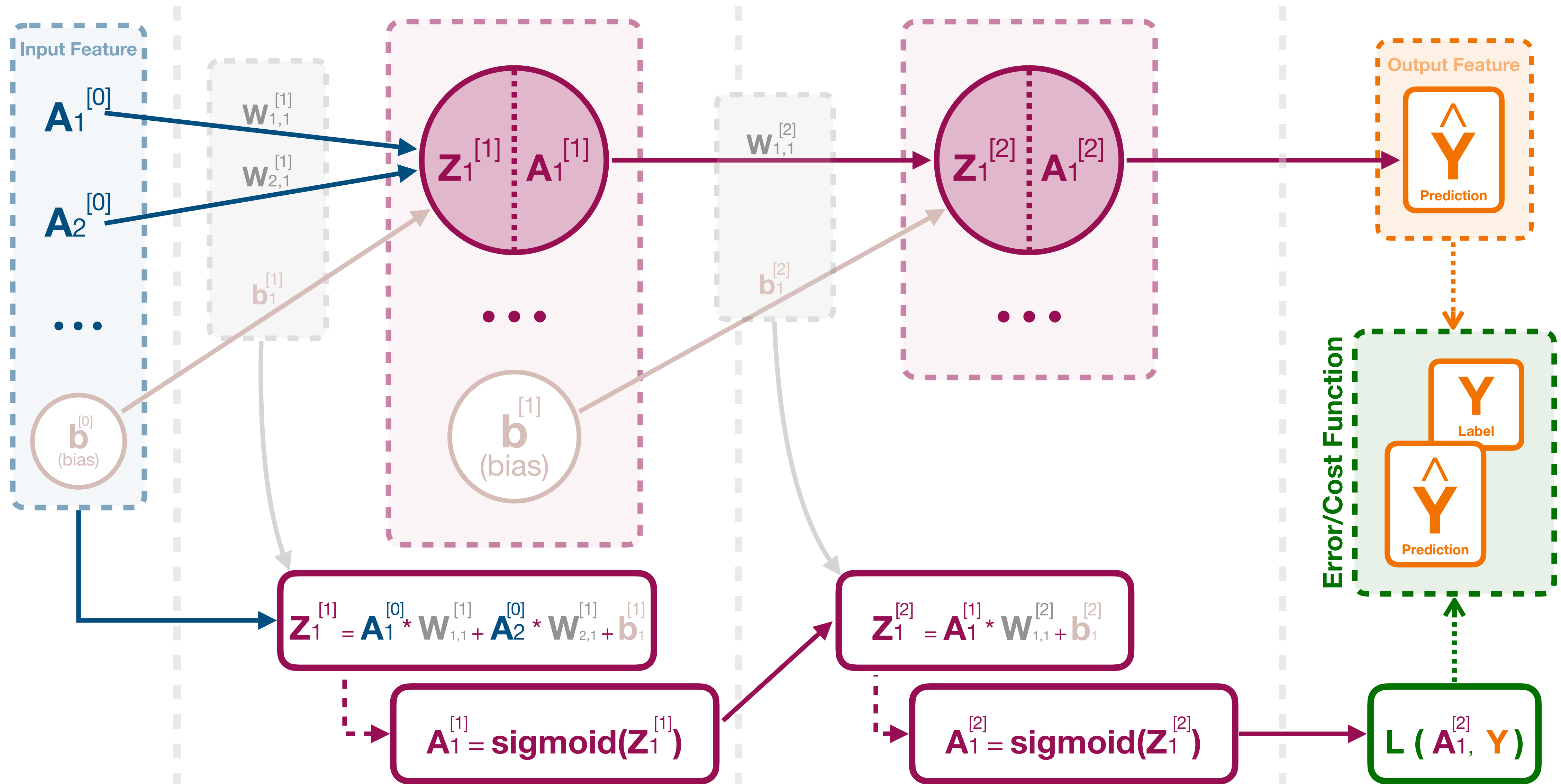


Computational Graph ($f(g(h(x)))$)



Chain Rule

Regla de la Cadena

Derivada de la Cost Function

$$L(A_1^{[2]}, Y)$$

Derivada: Aggregation Function

$$\partial Z$$

Formula

$$\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial A} * \frac{\partial A}{\partial Z} * \frac{\partial Z}{\partial W_1}$$

Calcular la Derivada del Valor de Error respecto al Valor de un Parámetro

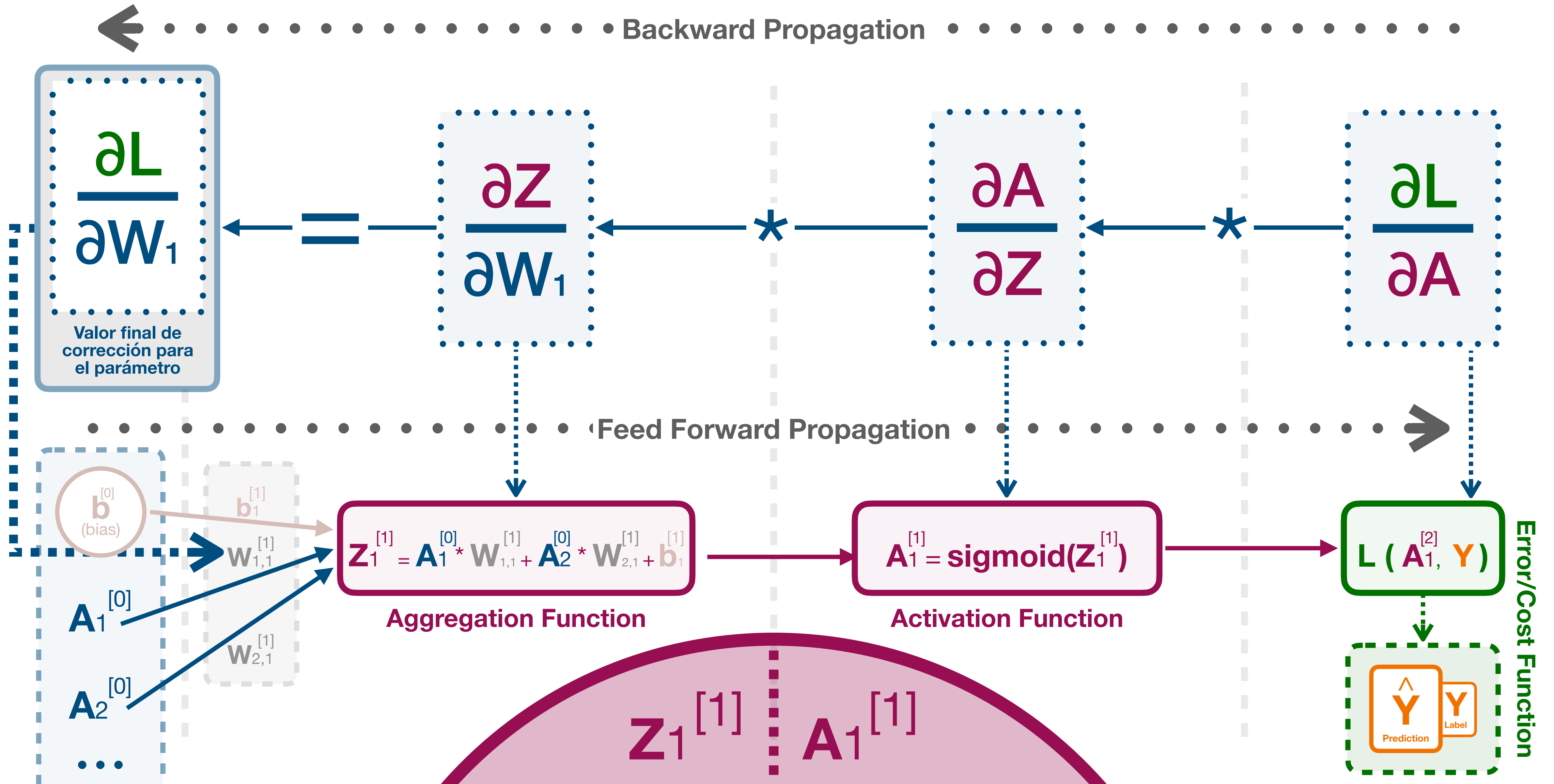
Derivada del parametro

$$W_{n,m}^{[k]}$$

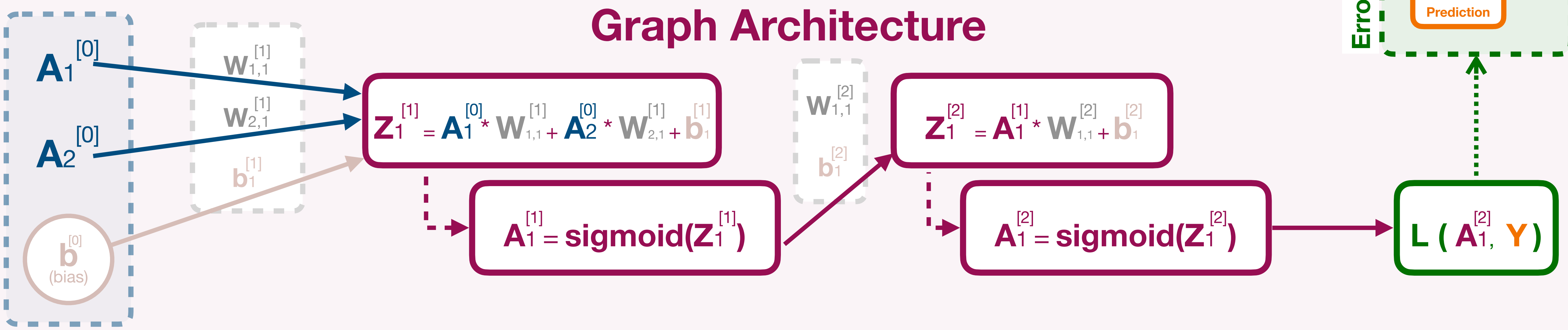
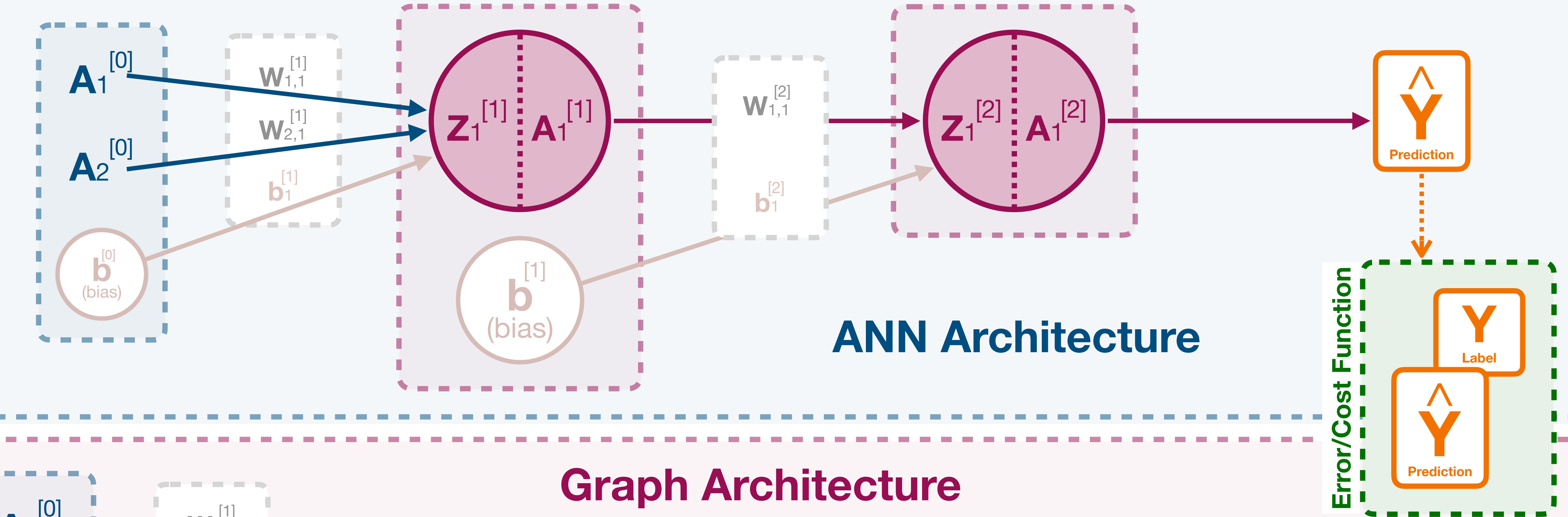
Derivada: Activation Function

$$\partial A$$

Derivatives with Computational Graph



Architectures



Derivatives Solutions

Resolución a partir del Grafo de la diapositiva previa

Backward Propagation

Backward Propagation

$$dw^{[2]} = a^{[1]} \cdot (a^{[2]} - y)$$

$$db^{[2]} = a^{[2]} - y$$

$$dw_1^{[1]} = (a^{[2]} - y) \cdot w^{[2]} \cdot (1 - a^{[1]2}) \cdot A_1^{[0]}$$

$$dw_2^{[1]} = (a^{[2]} - y) \cdot w^{[2]} \cdot (1 - a^{[1]2}) \cdot A_2^{[0]}$$

$$db^{[1]} = (a^{[2]} - y) \cdot w^{[2]} \cdot (1 - a^{[1]2})$$

Parameter update

$$w^{[2]} = w^{[2]} - \eta \cdot dw^{[2]}$$

$$b^{[2]} = b^{[2]} - \eta \cdot db^{[2]}$$

$$w_1^{[1]} = w_1^{[1]} - \eta \cdot dw_1^{[1]}$$

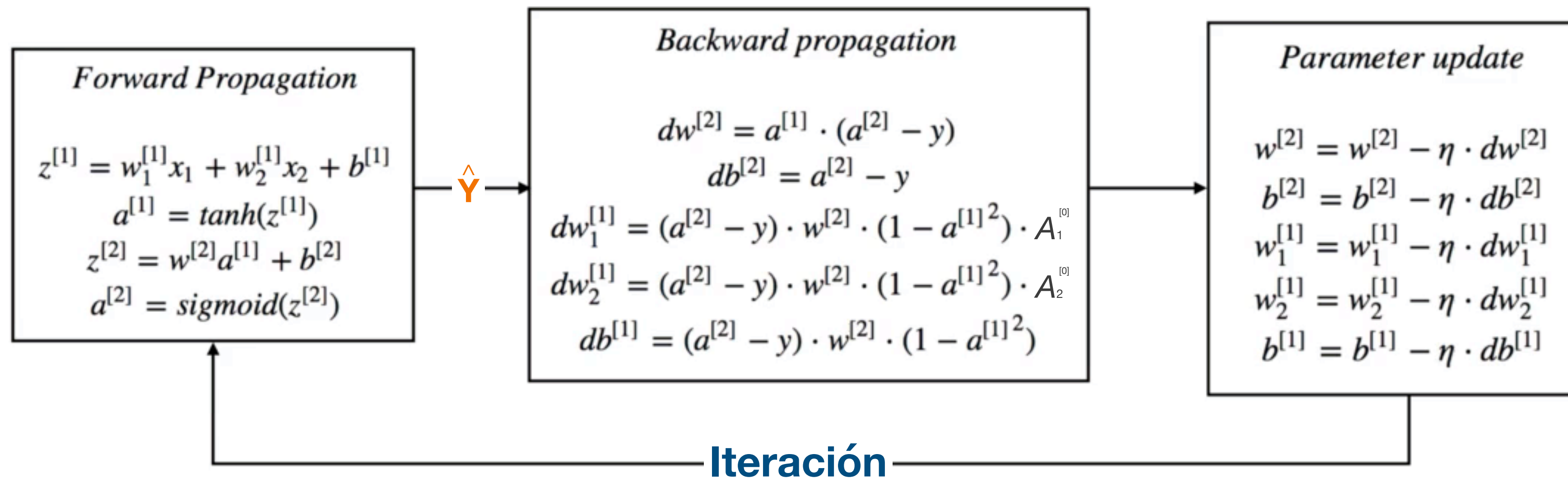
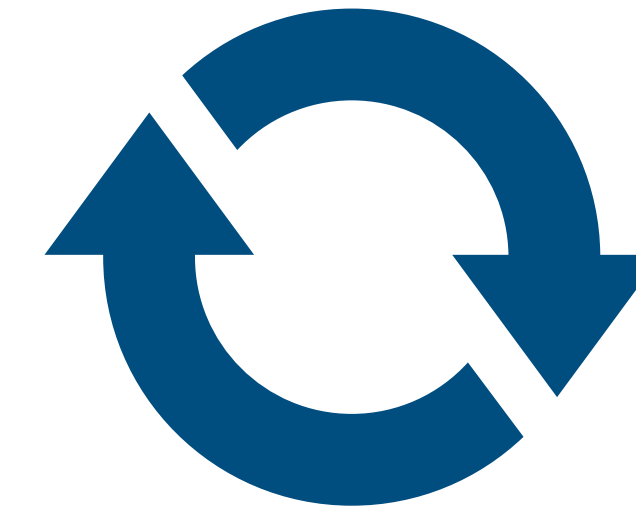
$$w_2^{[1]} = w_2^{[1]} - \eta \cdot dw_2^{[1]}$$

$$b^{[1]} = b^{[1]} - \eta \cdot db^{[1]}$$

ANN Training Process

Proceso de entrenamiento de una RNA

Entrenando RNAs profundas



Flujo hacia adelante del proceso de ejecución de la **RNA** a partir de un **vector de entrada** (conjunto de características) hasta producir el **valor de predicción**

Flujo hacia atrás determinando las **derivadas** de las **Funciones de Agregación y Activación** a partir de la **derivada** de la **Función de Coste/Error**

Actualización del valor de los parámetros de la **RNA** de manera individual a lo largo de **todas las capas**