딥러닝 올인원
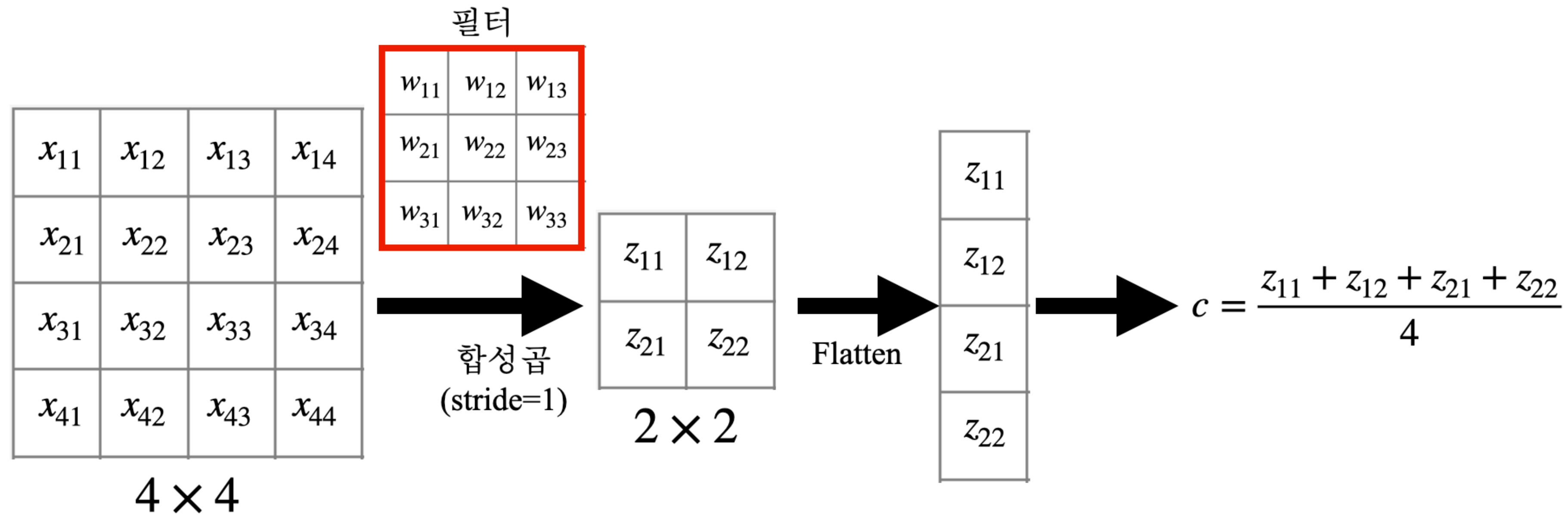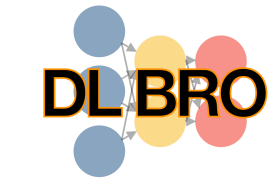
합성곱 신경망의 최적화
16강
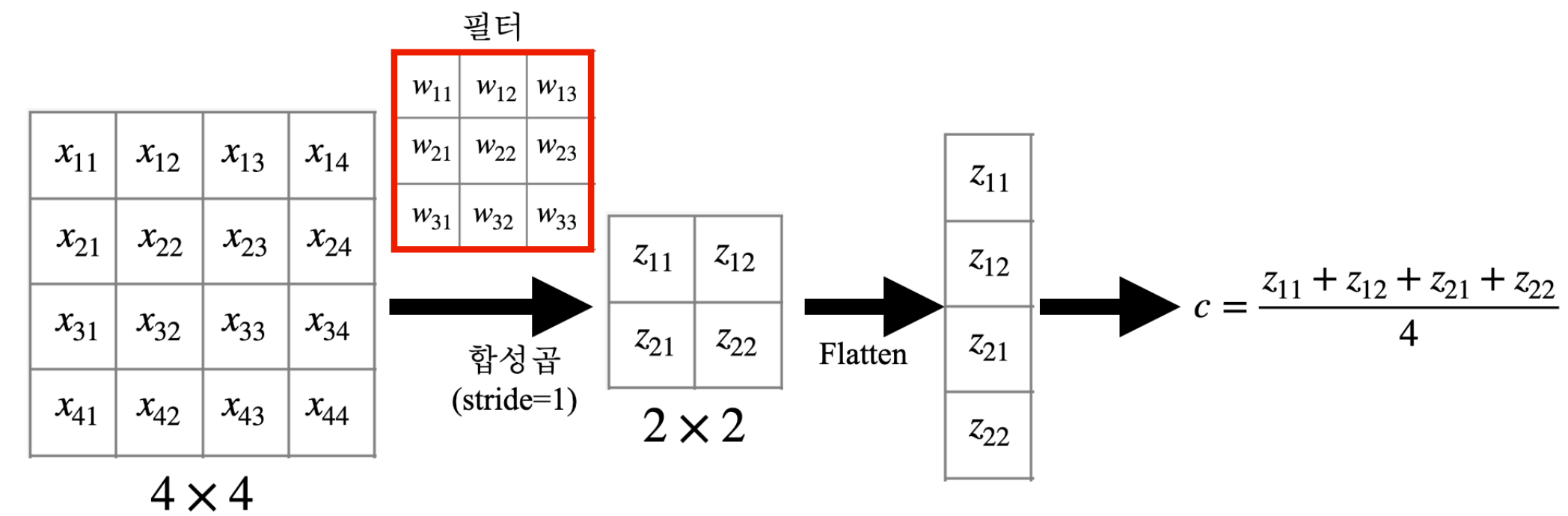
딥러닝호형

필터

$$\frac{\partial c}{\partial w_{11}} = \frac{\partial c}{\partial z_{11}}\frac{\partial z_{11}}{\partial w_{11}} + \frac{\partial c}{\partial z_{12}}\frac{\partial z_{12}}{\partial w_{11}} + \frac{\partial c}{\partial z_{21}}\frac{\partial z_{21}}{\partial w_{11}} + \frac{\partial c}{\partial z_{22}}\frac{\partial z_{22}}{\partial w_{11}} = \frac{1}{4}(\frac{\partial z_{11}}{\partial w_{11}} + \frac{\partial z_{12}}{\partial w_{11}} + \frac{\partial z_{21}}{\partial w_{11}} + \frac{\partial z_{22}}{\partial w_{11}})$$

필터

4 × 4 행렬 x와 3 × 3 필터 w의 합성곱(stride=1)으로 2 × 2 행렬 z를 구한 뒤 Flatten하여 $c = \dfrac{z_{11} + z_{12} + z_{21} + z_{22}}{4}$
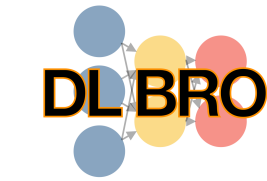
$$\frac{\partial c}{\partial w_{11}} = \frac{\partial c}{\partial z_{11}}\frac{\partial z_{11}}{\partial w_{11}} + \frac{\partial c}{\partial z_{12}}\frac{\partial z_{12}}{\partial w_{11}} + \frac{\partial c}{\partial z_{21}}\frac{\partial z_{21}}{\partial w_{11}} + \frac{\partial c}{\partial z_{22}}\frac{\partial z_{22}}{\partial w_{11}} = \frac{1}{4}\left(\frac{\partial z_{11}}{\partial w_{11}} + \frac{\partial z_{12}}{\partial w_{11}} + \frac{\partial z_{21}}{\partial w_{11}} + \frac{\partial z_{22}}{\partial w_{11}}\right)$$
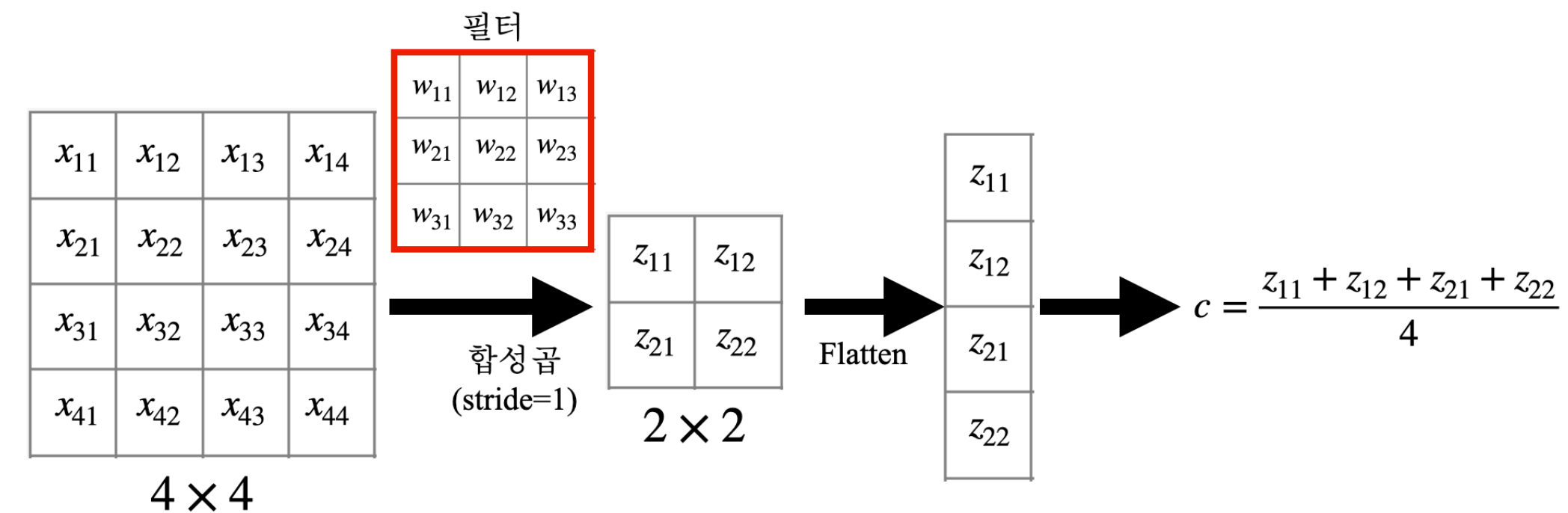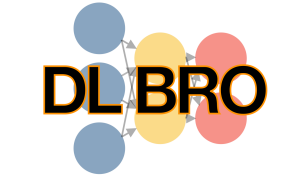
$$z_{11} = x_{11}w_{11} + x_{12}w_{12} + x_{13}w_{13} + x_{21}w_{21} + x_{22}w_{22} + x_{23}w_{23} + x_{31}w_{31} + x_{32}w_{32} + x_{33}w_{33}$$

$$z_{12} = x_{12}w_{11} + \dots \qquad z_{21} = x_{21}w_{11} + \dots \qquad z_{22} = x_{22}w_{11} + \dots$$

$$\frac{\partial z_{11}}{\partial w_{11}} = x_{11} \qquad \frac{\partial z_{12}}{\partial w_{11}} = x_{12} \qquad \frac{\partial z_{21}}{\partial w_{11}} = x_{21} \qquad \frac{\partial z_{22}}{\partial w_{11}} = x_{22}$$

$$\frac{\partial c}{\partial w_{11}} = \frac{1}{4}(x_{11} + x_{12} + x_{21} + x_{22}) \qquad w_{11} \leftarrow w_{11} - \alpha\frac{\partial c}{\partial w_{11}}$$

$$\frac{\partial c}{\partial w_{11}} = \frac{\partial c}{\partial z_{11}}\frac{\partial z_{11}}{\partial w_{11}} + \frac{\partial c}{\partial z_{12}}\frac{\partial z_{12}}{\partial w_{11}} + \frac{\partial c}{\partial z_{21}}\frac{\partial z_{21}}{\partial w_{11}} + \frac{\partial c}{\partial z_{22}}\frac{\partial z_{22}}{\partial w_{11}}$$

$$\frac{\partial c}{\partial w_{12}} = \frac{\partial c}{\partial z_{11}}\frac{\partial z_{11}}{\partial w_{12}} + \frac{\partial c}{\partial z_{12}}\frac{\partial z_{12}}{\partial w_{12}} + \frac{\partial c}{\partial z_{21}}\frac{\partial z_{21}}{\partial w_{12}} + \frac{\partial c}{\partial z_{22}}\frac{\partial z_{22}}{\partial w_{12}}$$

....

$$\frac{\partial c}{\partial w_{33}} = \frac{\partial c}{\partial z_{11}}\frac{\partial z_{11}}{\partial w_{33}} + \frac{\partial c}{\partial z_{12}}\frac{\partial z_{12}}{\partial w_{33}} + \frac{\partial c}{\partial z_{21}}\frac{\partial z_{21}}{\partial w_{33}} + \frac{\partial c}{\partial z_{22}}\frac{\partial z_{22}}{\partial w_{33}}$$

$$\nabla_w c = \begin{pmatrix} \dfrac{\partial c}{\partial w_{11}} \\ \dfrac{\partial c}{\partial w_{12}} \\ \vdots \\ \dfrac{\partial c}{\partial w_{33}} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial z_{11}}{\partial w_{11}} & \dfrac{\partial z_{12}}{\partial w_{11}} & \dfrac{\partial z_{21}}{\partial w_{11}} & \dfrac{\partial z_{22}}{\partial w_{11}} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial z_{11}}{\partial w_{33}} & \dfrac{\partial z_{12}}{\partial w_{33}} & \dfrac{\partial z_{21}}{\partial w_{33}} & \dfrac{\partial z_{22}}{\partial w_{33}} \end{pmatrix} \begin{pmatrix} \dfrac{\partial c}{\partial z_{11}} \\ \dfrac{\partial c}{\partial z_{12}} \\ \dfrac{\partial c}{\partial z_{21}} \\ \dfrac{\partial c}{\partial z_{22}} \end{pmatrix} \qquad w \leftarrow w - \alpha \nabla_w c$$