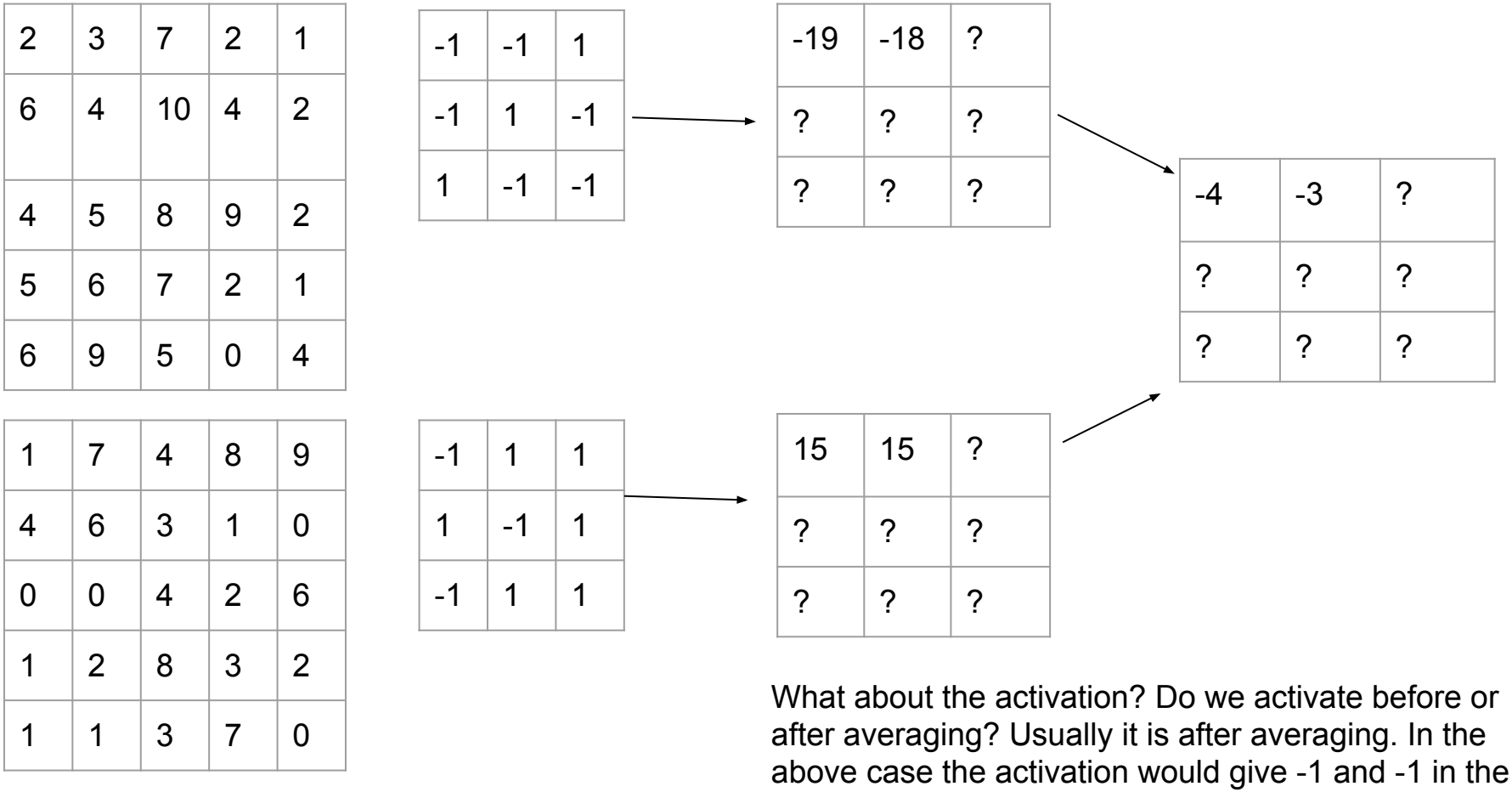
An RGB image has three sets of pixel values whereas a black and white has just one. The pixel values are between 0 and 255 in both cases. For RGB images we call each set a channel. How do we apply a convolution to an image with three channels? Below we see on the left an image with two channels. We apply a convolutional kernel with two input channels to the image and average the output.



What about the activation? Do we activate before or after averaging? Usually it is after averaging. In the above case the activation would give -1 and -1 in the final output of the convolution,.