softmax invariance to offsets by a constant

(a) (5 points) Prove that softmax is invariant to constant offsets in the input, that is, for any input vector x and any constant c,

$$\text{softmax}(x) = \text{softmax}(x + c)$$

where $x + c$ means adding the constant c to every dimension of x. Remember that

$$softmax(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \tag{1}$$

Note: In practice, we make use of this property and choose $c = -\max_i x_i$ when computing softmax probabilities for numerical stability (i.e., subtracting its maximum element from all elements of x).

We know

$$\text{softmax}(x + c)_i = \frac{e^{x_i + c}}{\sum_j e^{x_j + c}} = \frac{e^{x_j} e^c}{\sum_j e^{x_j} e^c} = \frac{e^{x_i}}{\sum_j e^{x_j}}$$

since (e^c) cancel out.