$$y = \gamma \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta$$

μ = batch mean
σ^2= batch variance
ϵ = small constant to avoid /0 errors
γ, β = learned parameters to restore non-linear behaviour (after the activation function)

Substituting in the convolution expression:

$$y = \gamma \frac{(Wx + B) - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta$$

With a little rearranging:

$$y = \frac{\gamma W}{\sqrt{\sigma^2 + \epsilon}} x + \gamma \frac{B - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta$$

Now we get the update equations for W and B:

$$W \rightarrow \frac{\gamma W}{\sqrt{\sigma^2 + \epsilon}}$$

$$B \rightarrow \gamma \frac{B - \mu}{\sqrt{\sigma^2 + \epsilon}} + \beta$$