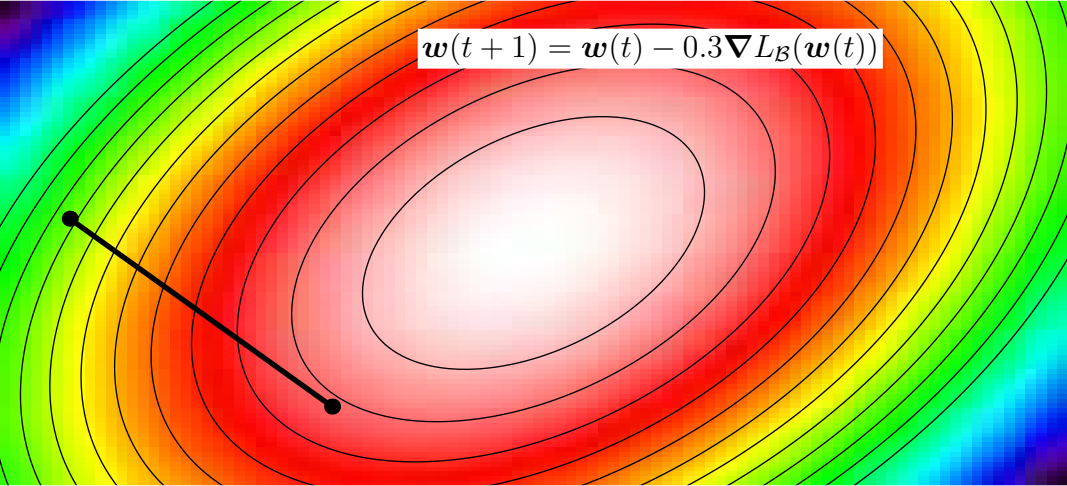
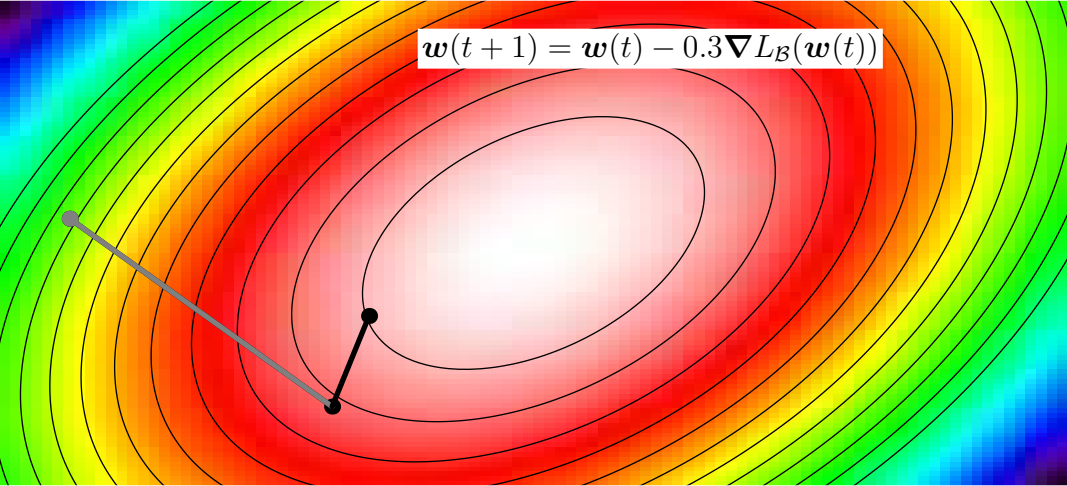


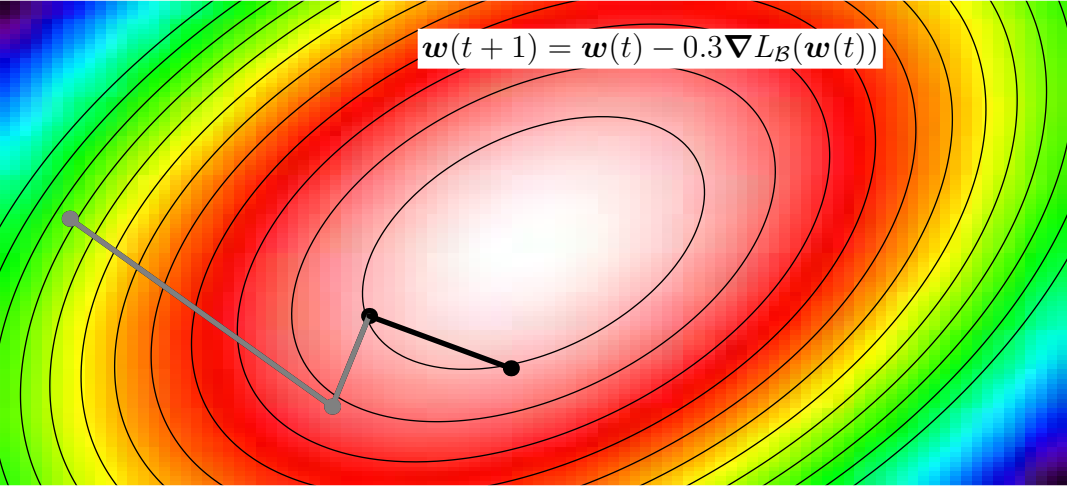
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



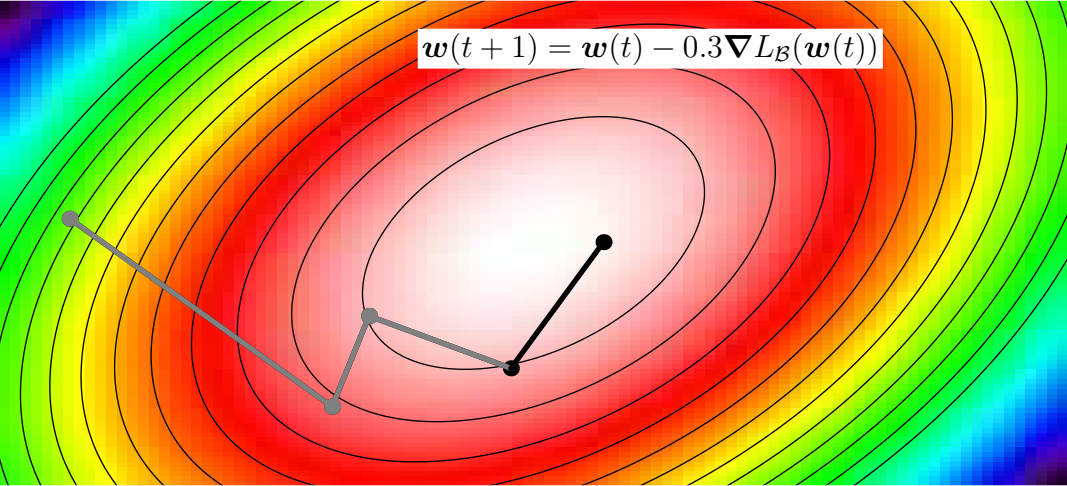
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



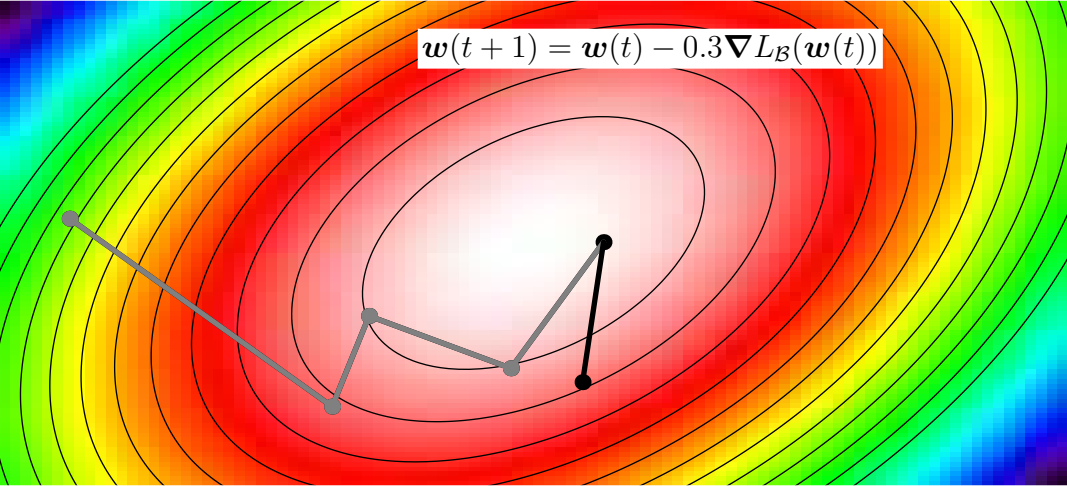
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



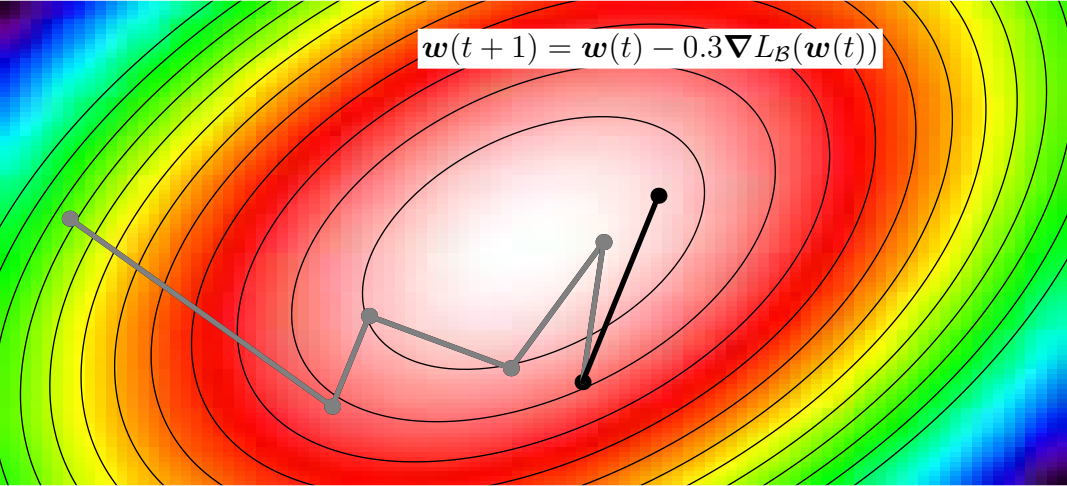
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



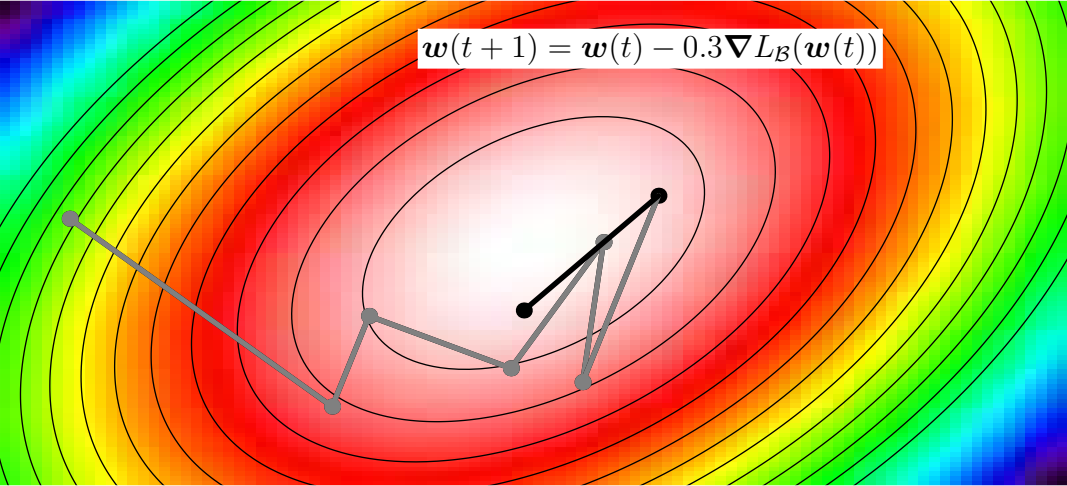
$$\mathbf{w}(t+1) = \mathbf{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\mathbf{w}(t))$$



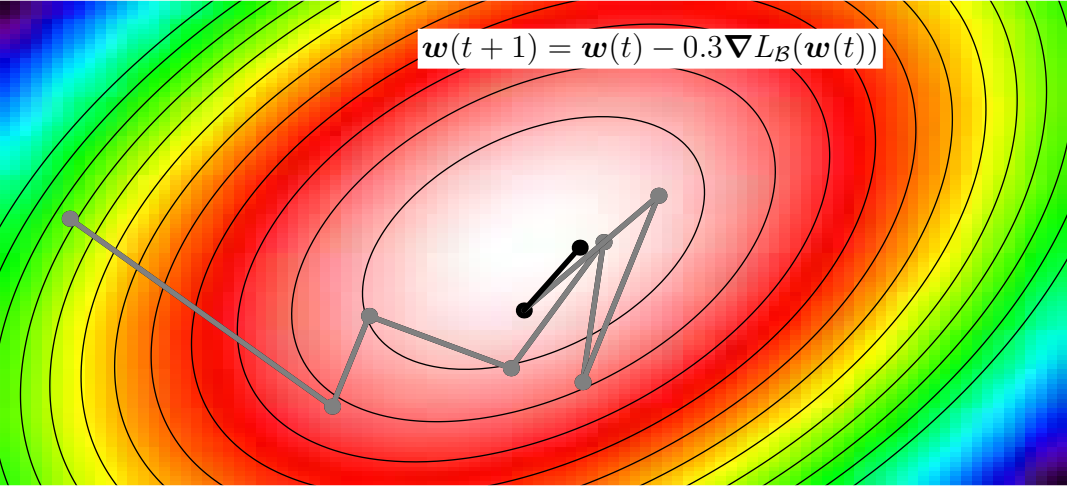
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



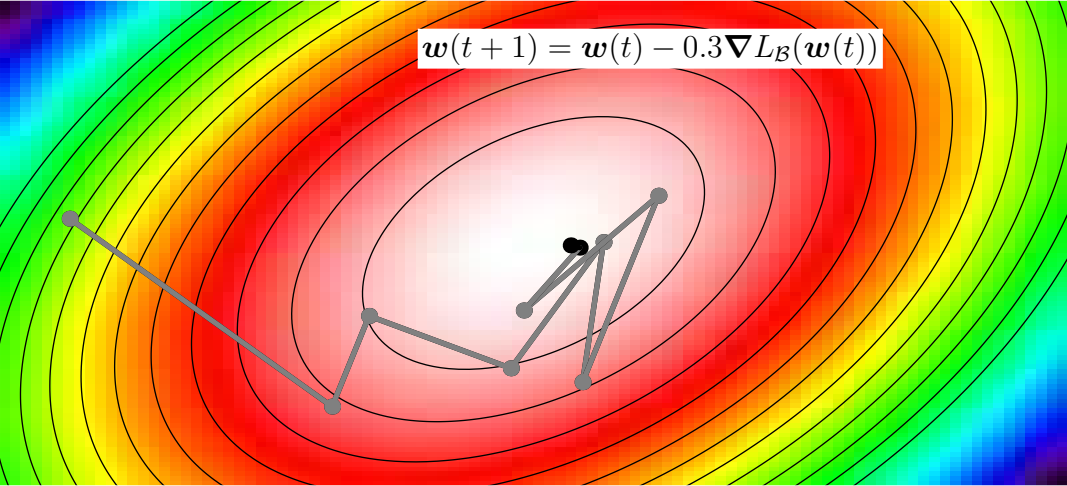
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



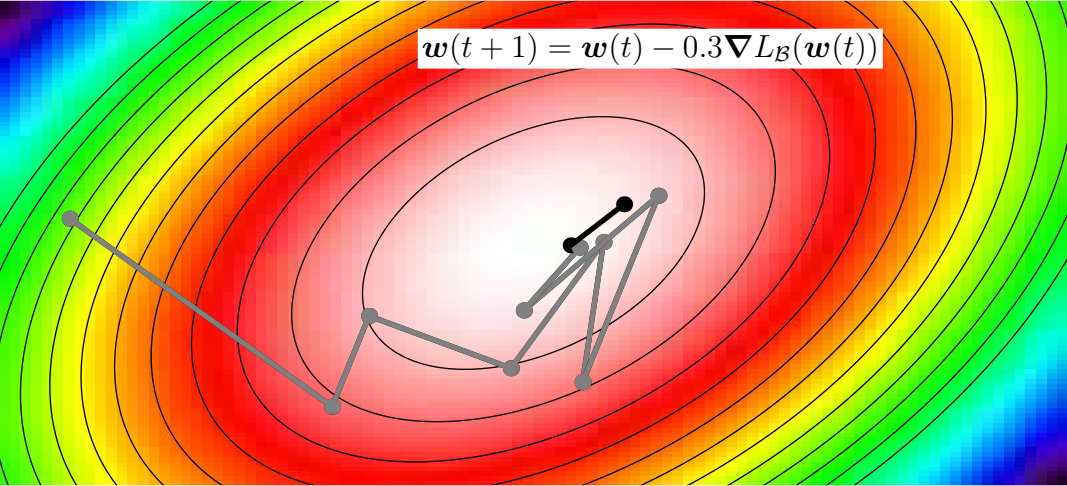
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



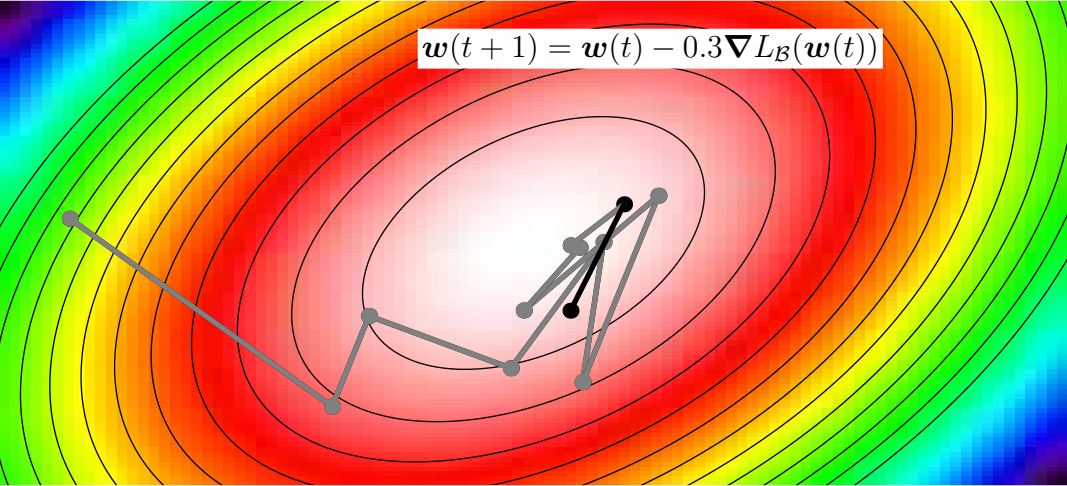
$$\mathbf{w}(t+1) = \mathbf{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\mathbf{w}(t))$$



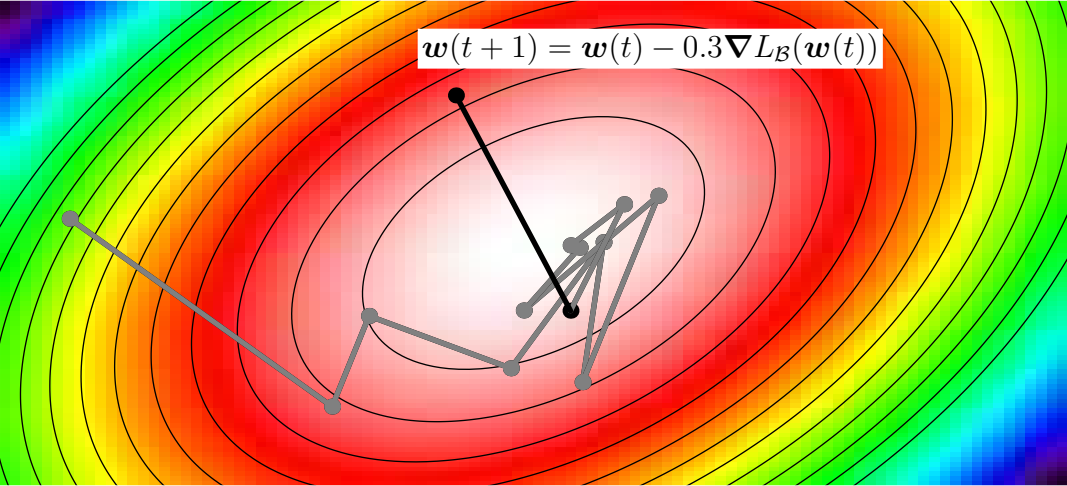
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



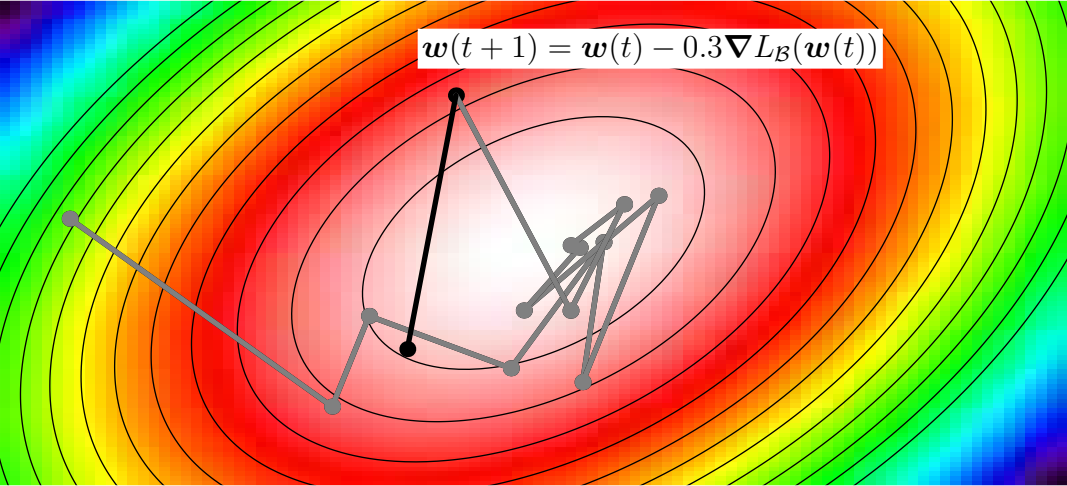
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



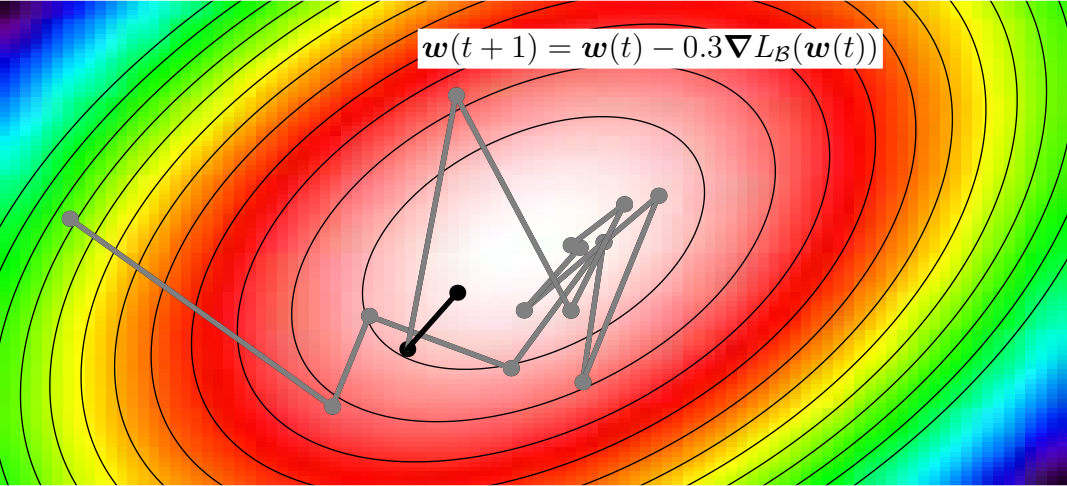
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



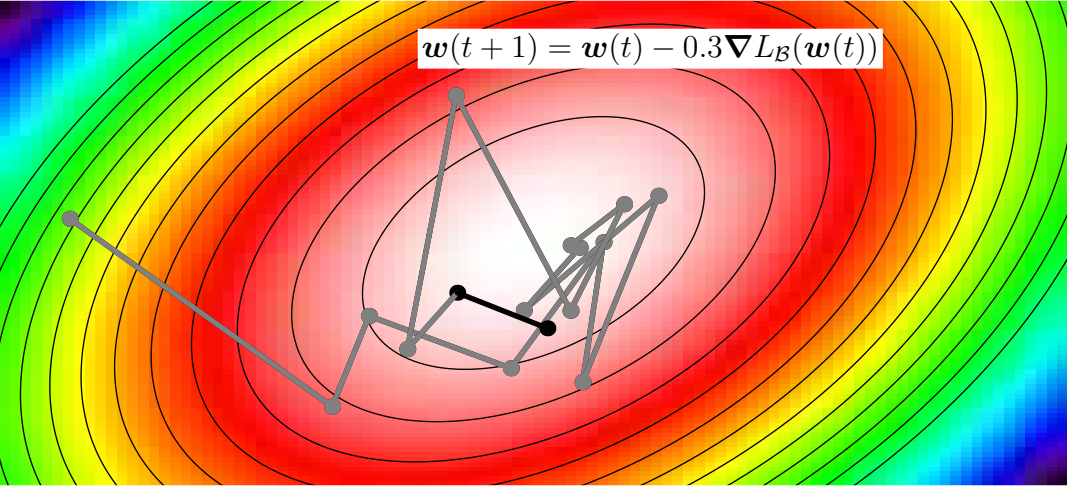
$$\mathbf{w}(t+1) = \mathbf{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\mathbf{w}(t))$$

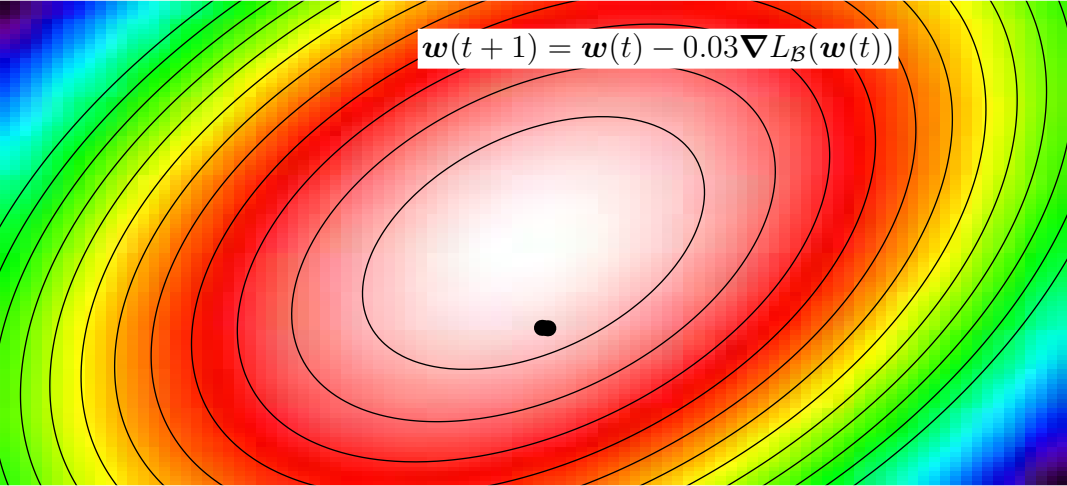


$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.3 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$

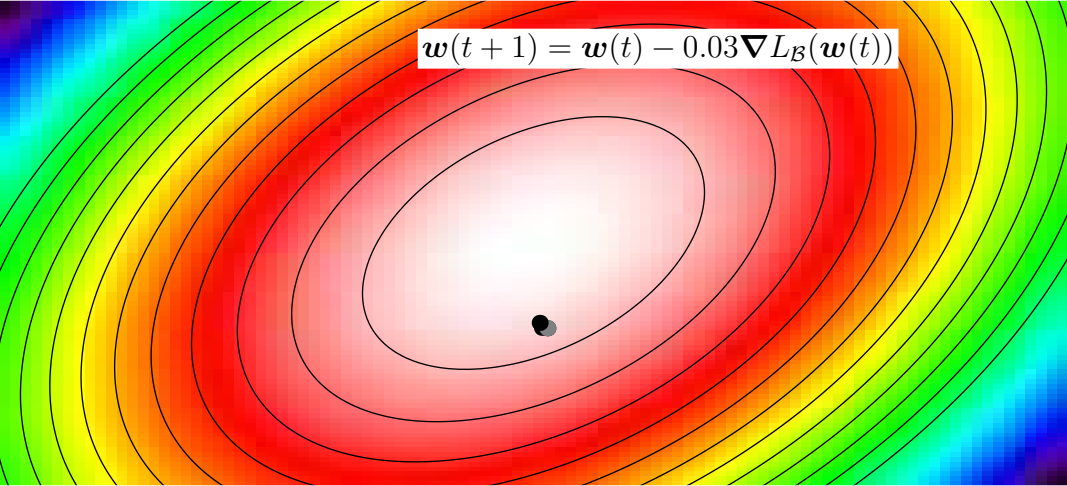




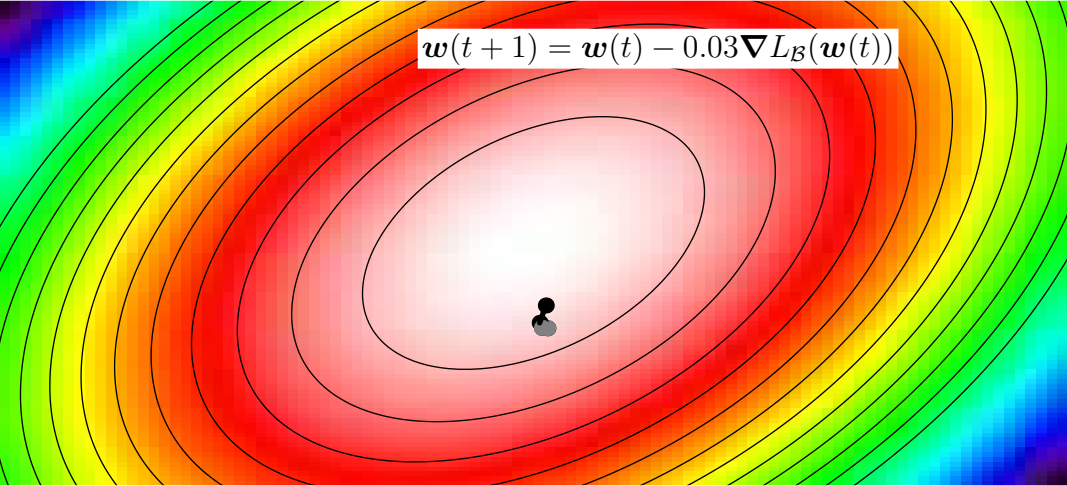
A contour plot of a loss function $L_{\mathcal{B}}$ is shown. The plot features concentric elliptical contours, with the center being the minimum of the function. The contours are colored using a rainbow gradient, where red and white represent the lowest values (the minimum) and blue and green represent higher values. A black dot is placed on one of the inner contours, representing the current parameter vector $w(t)$. A white rectangular box is overlaid on the upper part of the plot, containing the update equation for the next step $w(t+1)$.

$$w(t+1) = w(t) - 0.03 \nabla L_{\mathcal{B}}(w(t))$$

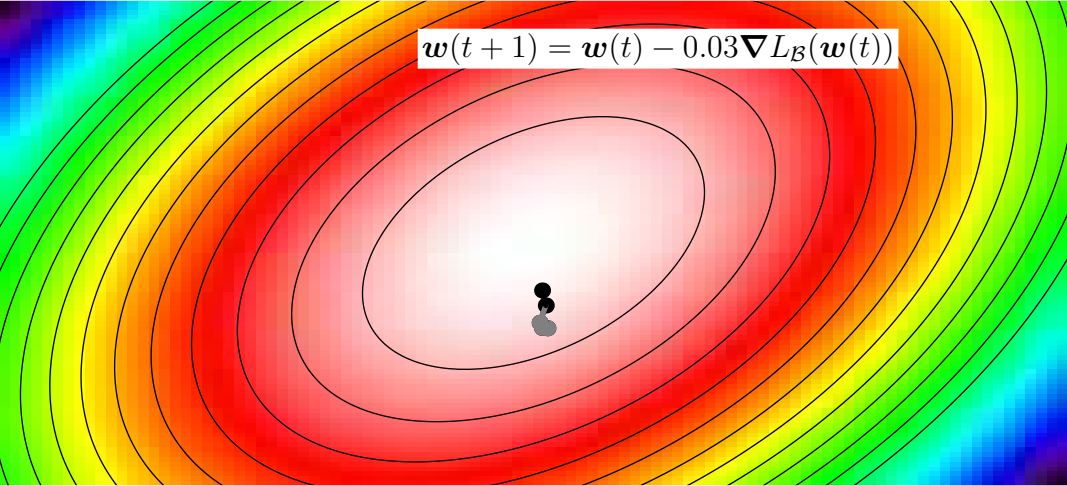
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



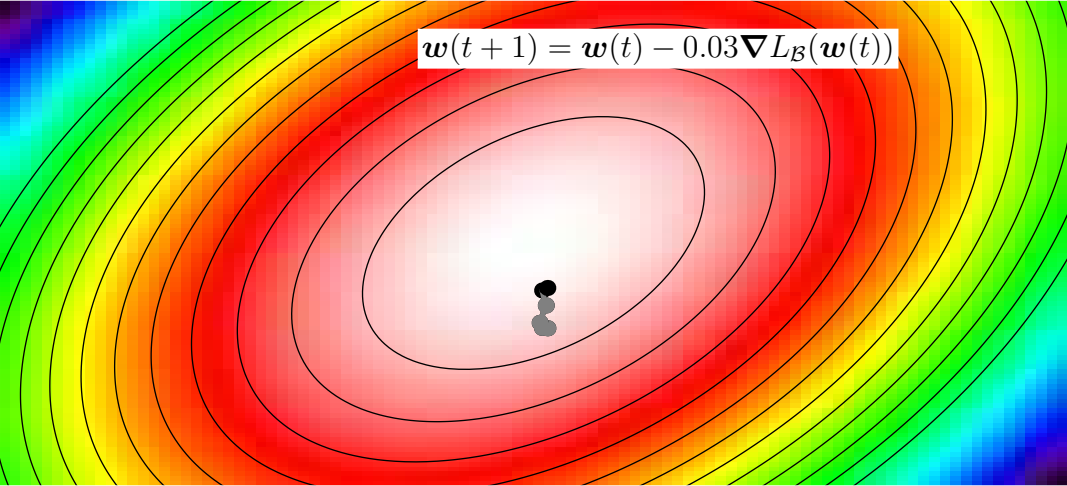
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



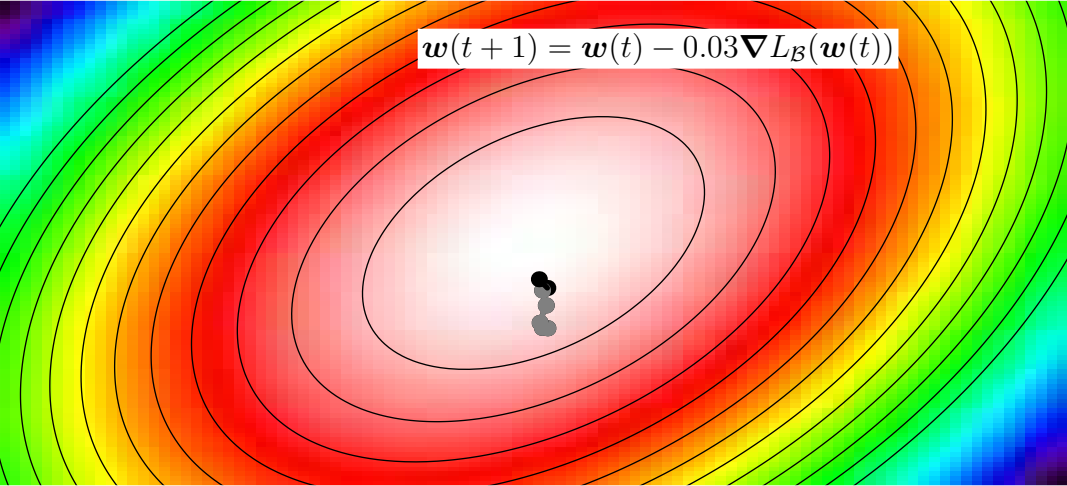
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



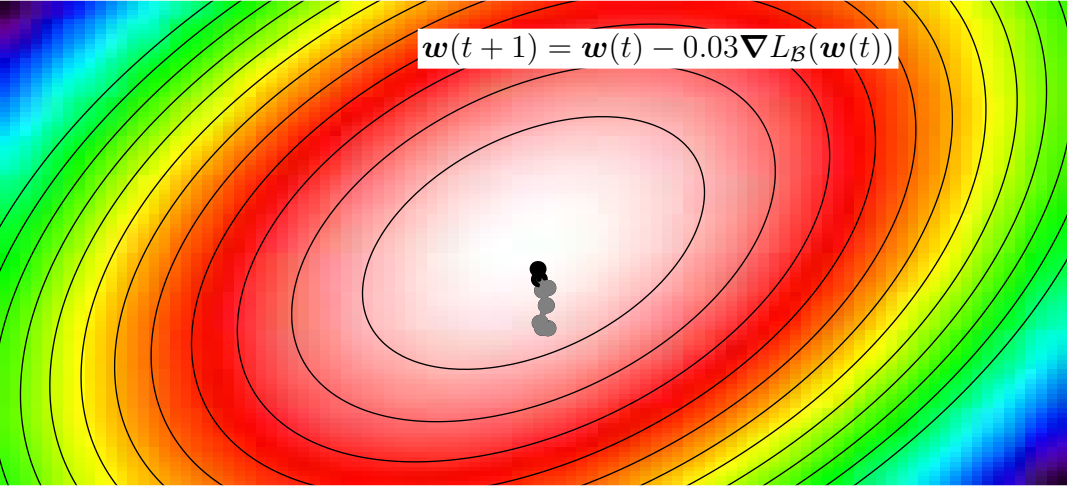
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



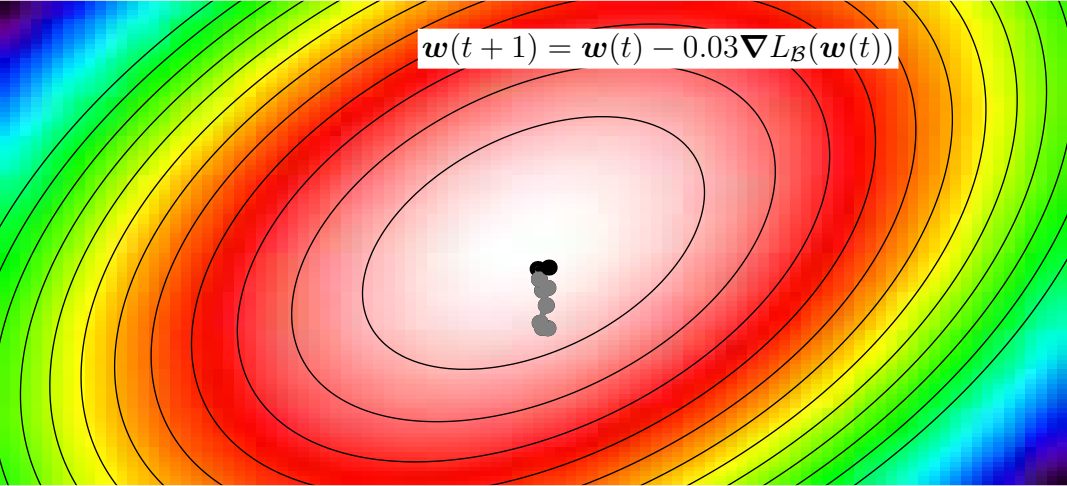
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



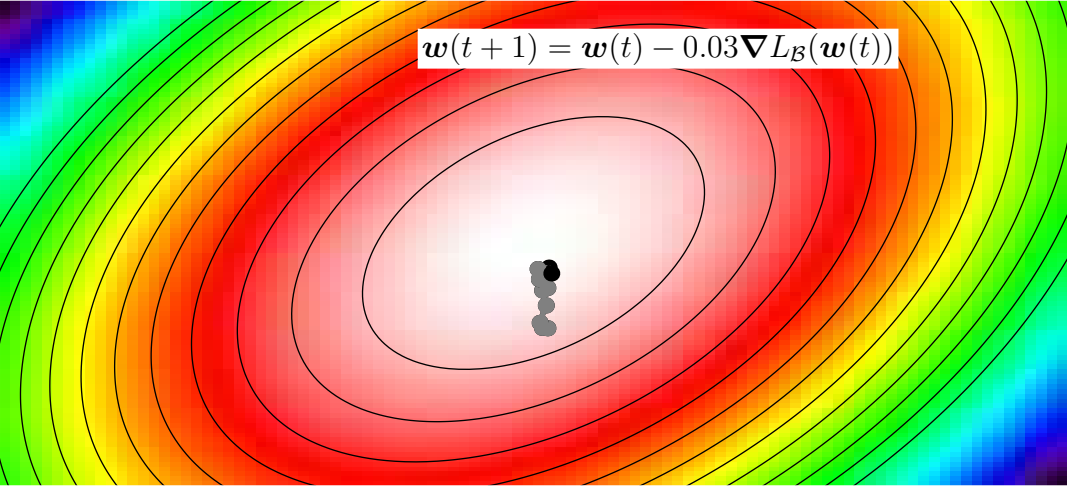
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



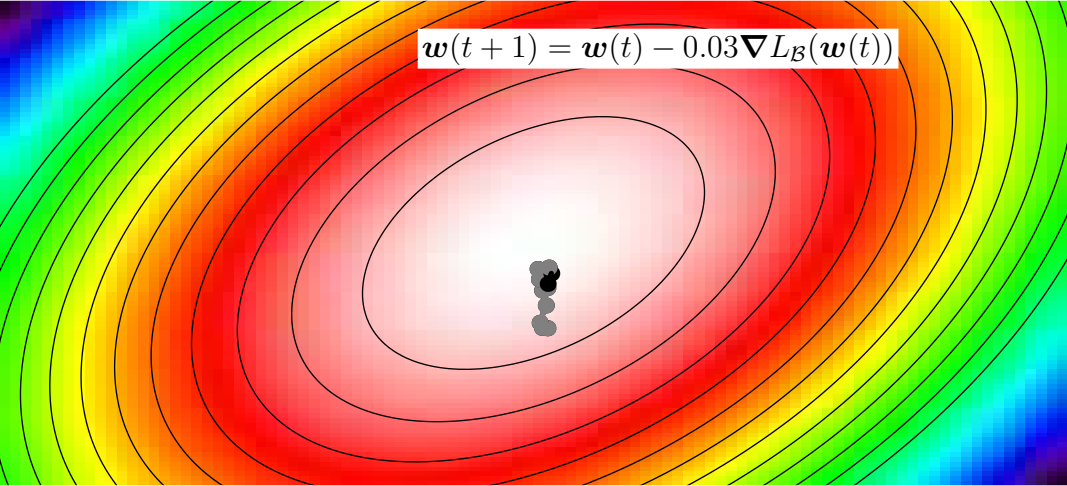
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



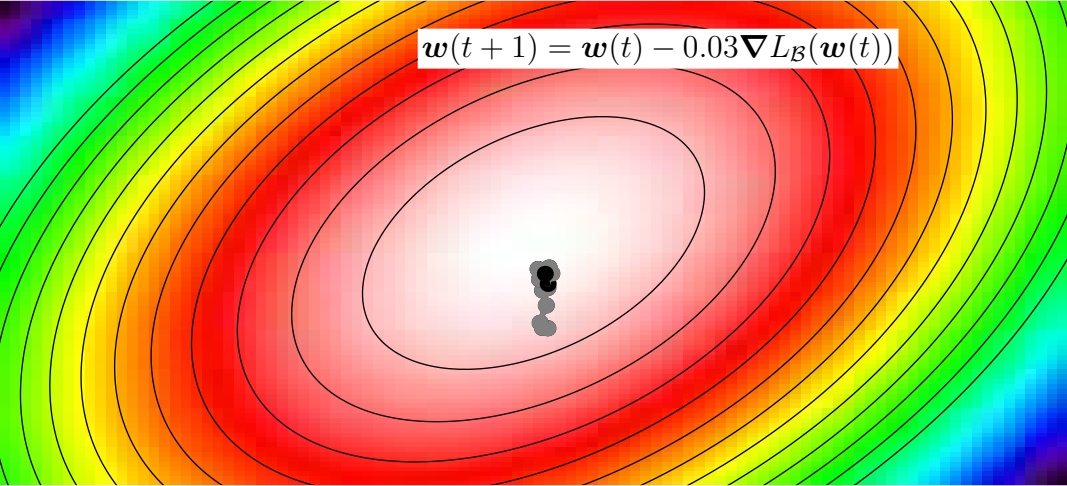
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



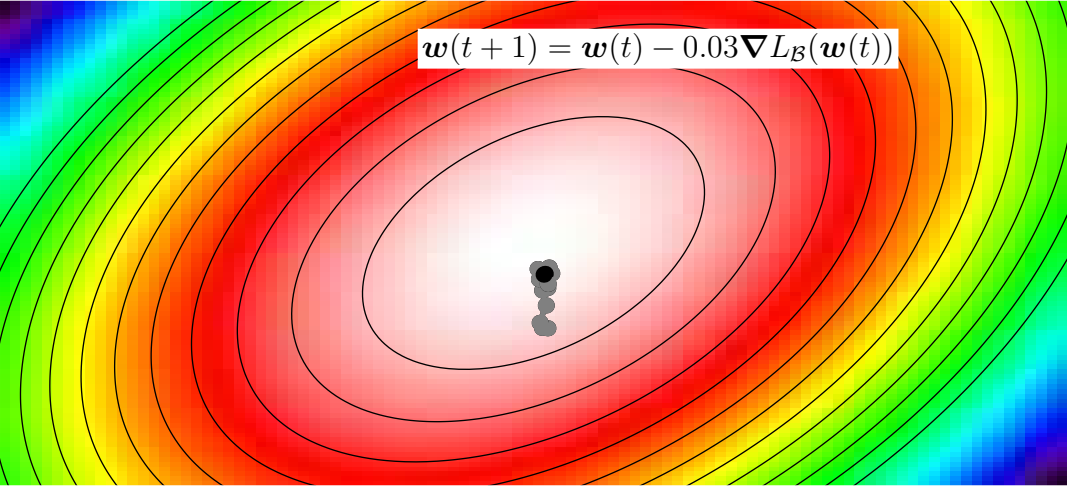
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



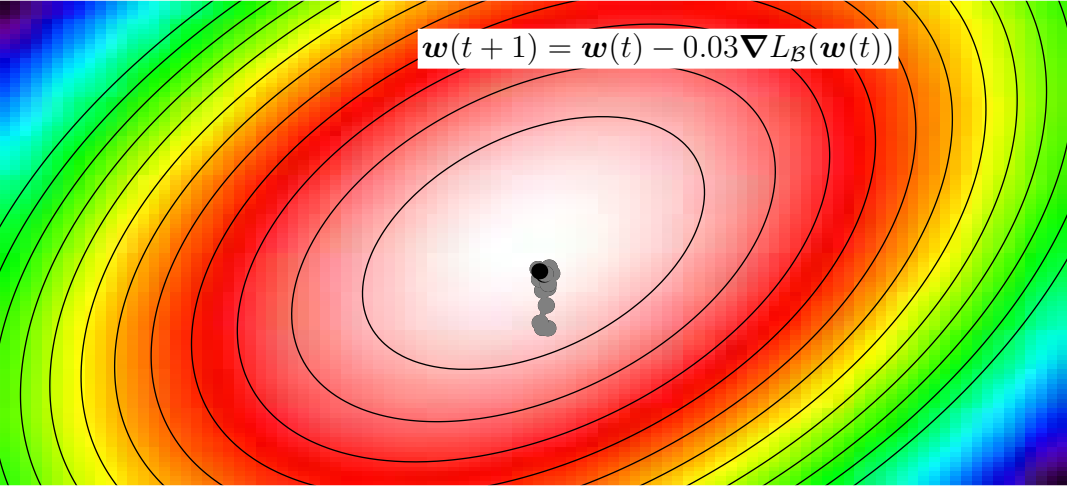
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



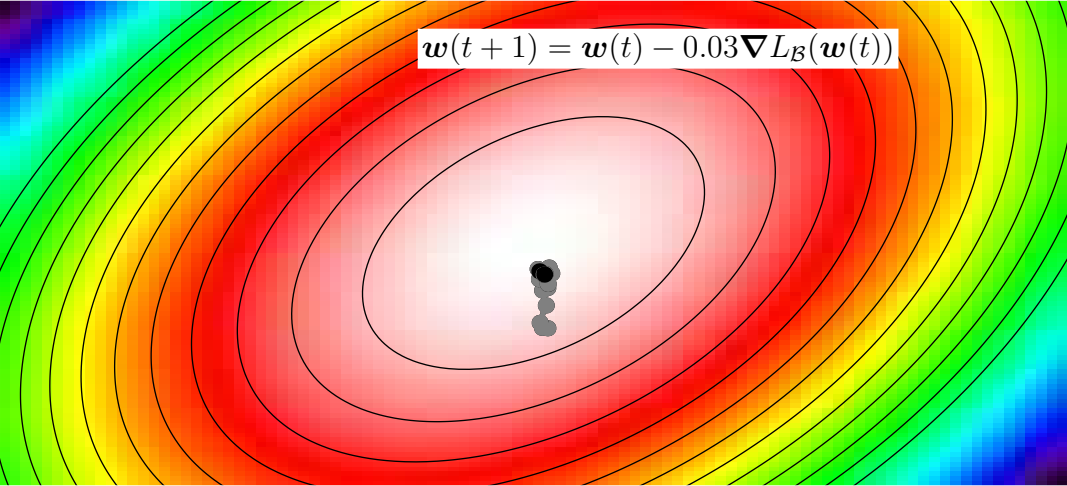
$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$



$$\boldsymbol{w}(t+1) = \boldsymbol{w}(t) - 0.03 \nabla L_{\mathcal{B}}(\boldsymbol{w}(t))$$

