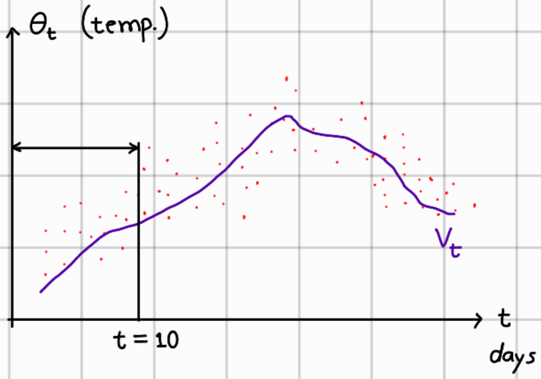
PHY654
Machine Learning in Particle Physics

Saturday, 19 August, 2024

Exponentially Weighted Average (EWA) or Moving Average

$$V_t = \beta V_{t-1} + (1-\beta)\theta_t \quad , \quad \beta = 0.9 \text{ is usually a popular choice.}$$



Bias correction

$$V_t^{corr.} = \frac{V_t}{1-\beta^t} \approx V_t$$

$\beta = 0.9$, $\theta_1 = 40^\circ$ (A.U.), $V_0 = 0$,

$V_1 = 0.1 * 40^\circ$

$V_2 = \dots$

$\vdots$

V_{10}

1. Gradient Descent with Momentum

on iter t

$$\frac{\partial \mathcal{L}}{\partial w} \equiv dW, \qquad \frac{\partial \mathcal{L}}{\partial b} \equiv db$$

compute dW, db

Forward propagation: $X \longrightarrow \square \longrightarrow \square \longrightarrow \hat{y}$ $\quad \mathcal{L}$

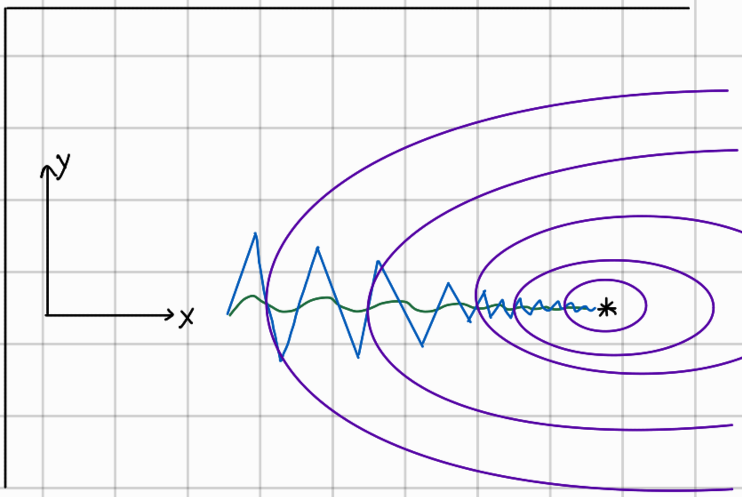
Backward propagation: $\longleftarrow \square \longleftarrow \square \longleftarrow y$

$$V_{dW}^{(t)} = \beta V_{dW}^{(t-1)} + (1-\beta)\, dW^{(t)}$$

$$V_{db}^{(t)} = \beta V_{db}^{(t-1)} + (1-\beta)\, db^{(t)}$$

Update Rule:

$$\boxed{\begin{aligned} W &:= W - \alpha V_{dW} \\ b &:= b - \alpha V_{db} \end{aligned}}$$



2. RMS Propagation:

$$S_{dW}^{(t)} = \beta S_{dW}^{(t-1)} + (1-\beta) dW^{2\,(t)}$$

$$S_{db}^{(t)} = \beta S_{db}^{(t-1)} + (1-\beta) db^{2(t)}$$

Update rule:

$$W := W - \alpha \frac{dW}{\sqrt{S_{dW} + \varepsilon}}$$

$$b := b - \alpha \frac{db}{\sqrt{S_{db} + \varepsilon}}$$

ADAM optimization (Adaptive moment Estimation) [t: # of iter.]

$$\left.\begin{aligned}
V_{dW}^{(t)} &= \beta_1 V_{dW}^{(t-1)} + (1-\beta_1) dW^{(t)} \\
V_{du}^{(t)} &= \beta_1 V_{du}^{(t-1)} + (1-\beta_1) du^{(t)} \\
S_{dW}^{(t)} &= \beta_2 S_{dW}^{(t-1)} + (1-\beta_2) dW^{2\,(t)} \\
S_{du}^{(t)} &= \beta_2 S_{du}^{(t-1)} + (1-\beta_2) du^{2\,(t)}
\end{aligned}\right\} \beta_1 = 0.9,\ \beta_2 = 0.999,\ \varepsilon = 10^{-8}$$

$$V_{dW}^{corr.} = \frac{V_{dW}}{1-\beta_1^t}$$

$$V_{du}^{corr.} = \frac{V_{du}}{1-\beta_1^t}$$

$$S_{dW}^{corr.} = \frac{S_{dW}}{1-\beta_2^t}$$

$$S_{du}^{corr.} = \frac{S_{du}}{1-\beta_2^t}$$

Update rule:

$$W := W - \alpha \frac{V_{dW}^{corr.}}{\sqrt{S_{dW}^{corr.} + \varepsilon}}$$

$$b := b - \alpha \frac{V_{du}^{corr.}}{\sqrt{S_{du}^{corr.} + \varepsilon}}$$

Wave Equation: (a 2nd order PDE)

$$\boxed{\frac{\partial^2 g(x,t)}{\partial t^2} = c^2 \frac{\partial^2 g(x,t)}{\partial x^2}}$$

Analytical solution:

$$g(t,x) = \sin(\pi x)\cot(\pi t) - \sin(\pi x)\sin(\pi t)$$

with boundary conditions:

$$\left.\begin{aligned}
& g(0,t) = g(1,t) = 0\ ;\ t \geq 0 \\
& g(x,0) = u(x)\ ;\ \left.\frac{\partial g(x,t)}{\partial t}\right|_{t=0} = v(x)
\end{aligned}\right\} x \in [0,1]$$

Trial Solution

$$g_t(x,t) = \underbrace{h_1(x,t)}_{} + x(1-x)t^2 N(x,t,p)$$

$$\Downarrow$$

$$h_1(x,t) = (1-t^2)u(x) + t\,v(x)$$

$$g_t(x,0) \overset{\checkmark}{=} u(x), \quad \left.\frac{\partial g_t(x,t)}{\partial t}\right|_{t=0} \overset{\checkmark}{=} v(x)$$

$$g_t(0,t) = (1-t^2)u(x) + t\,v(x)$$

$$g_t(1,t) = (1-t^2)u(x) + t\,v(x)$$

$u(x) = \sin(\pi x)$	$v(x) = -\pi\sin(\pi x)$	$c = 1$

Try :

$$\frac{\partial g(x,t)}{\partial t} = \frac{\partial^2 g(x,t)}{\partial x^2} \longrightarrow \text{(diffusion equation)}$$

with $g(0,t) = 0 = g(1,t)$; $t \geq 0$

$g(x,0) = u(x)$; $x \in [0,1]$

Analytical solution: $g(x,t) = \exp(-\pi^2 t)\sin(\pi x)$

neural network for solving differential eqn
physics arxiv

PINN – google search.

Mini batch gradient descent.