- Mouentum

Ada Grad > Ada M 6D -> 560 Vo 2 (0, 6x, y3) prezeg: $2(0, 1x, y3) \Rightarrow \sqrt{2(0, 1x, y3)}$ $\theta_{o} \sim (?)$ • $m_{t} = (1 - \beta_{t}) \nabla_{\theta} Q_{t-1} + \beta_{t} m_{t-1}$ Ams Grad • $S_{2} = (1 - \beta_{2}) \left(\nabla_{0} d_{1} + \beta_{2} S_{1} - \beta_{2} \right)$ S= max (S, S) $\bullet \quad \theta_{t} = \theta_{t-1} - \theta_{t} \frac{m_{t}}{\sqrt{s_{t} + \varepsilon}}$











