Optimization

problems
- Local optima
- Plateau
- Saddle point

Solve
- Newton's method
- Momentum

Momentum : $\theta_{k+1} = \theta_k - \alpha g_k$ $\quad g_k = \nabla_\theta L(\theta_k) + \mu g_{k-1}$

"blend-in" previous direction