

Top Equations in Deep Learning

February 15, 2020

Top Equations in Deep Learning

Equation		Use	Citation
1 Stochastic Gradient Descent	$\theta_{i+1} \leftarrow \theta_i - \alpha \sum_{x \in B} \frac{\partial L(x)}{\partial \theta}$	Optimization	
2 Backpropagation	$\frac{\partial L}{\partial \theta} = \frac{\partial L}{\partial z} \frac{\partial z}{\partial \theta}$	Computing Gradients	
3 Linear Layer	$y = Wx + b$	Base Layer in Deep Learning	
4 Convolutions	$(f * g)[n] = \sum_{m=0}^K f[n-m]g[m]$	Image Networks	
5 ReLU	$\text{ReLU}(x) = \max(0, x)$	Activation Function	
6 Recurrent Cells	$h_{t+1} = \tanh(W_{ih}x_t + b_{ih} + W_{hh}h_t + b_{hh})$	Sequence Modeling	
7 Self Attention	$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$	Transformers	
8 Generative Adversarial Netwrks	$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))]$	Generative Modeling	
9 REINFORCE	$\triangle \theta = \alpha r \frac{\partial \log p(a \pi^\theta(s))}{\partial \theta}$	Reinforcement Learning	
10 (Variational) Auto-Encoder		Likelihood Modeling	
11 Categorical Cross Entropy	$L(x, y) = -x_y + \log\left(\sum_j \exp(x_j)\right)$	Multiclass Classification	
12 Bellman Equation	$V(s_t) = \max_{a_t} \left(R(s_t, a_t) + \gamma \sum_{s_{t+1}} p(s_t, a_t, s_{t+1}) V(s_{t+1}) \right)$ $Q(s_t, a_t) = Q(s_t, a_t) + \alpha \cdot \left(R(s_t, a_t) + \gamma \cdot \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right)$	Reinforcement Learning	
13 Reparametrization Trick	$\mathbb{E}_{z \sim p(z \theta)} [f(z)] = \mathbb{E}_{\epsilon \sim p(\epsilon)} [f(g(\epsilon, \theta))]$	Differential Sampling	
14 ℓ_p Norm	$\ x\ _p = \left(\sum_{i=1}^n x_i ^p \right)^{1/p}$ $\ A\ _F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij} ^2}$	Distance, Regression Loss	
15 Batch Normalization	$y = \frac{x - \mathbb{E}[x]}{\sqrt{\text{Var}[x] + \epsilon}} * \gamma + \beta$	Regularization	
16 Mutual Information			

Table 1