Lecture 13

Huang Ziheng

December 19, 2017

1 Deep Learning

Big question: why CNN matches well for image data.
One hidden layer

Universal Approximator Thm On a compact set $\Omega \in \mathcal{R}^d$,$\forall f \in C^1$, $\forall \epsilon > 0$,, $\exists$ one hidden layer NN g, s.t. $\|f - g\|_\Omega \leq \epsilon$
So the approximation is proved, but convergence efficiency is bad, thus more than one layer.

Depth/Width-bounded networks $Inputdim = d$, $Width \leq d + 1$
?$\exists const$ c,$\forall f \in C^1, \forall \epsilon > 0, \exists g$ depth/width-bounded and $|weights| \leq c$, s.t. $\|f - g\|_\Omega \leq \epsilon$
Obviously we need some other conditions.
①Cybenko 1989, Lu et d "The Expressive Power of Neural Networks, A View from the Width", N2PS, 2017

2 Reinforce Learning

MDP, parameters$(\mathcal{P}, \mathcal{R})$ known, then to evaluate policy π by Bellman Expectation Operator we can find the solution.
To find the optimal policy, policy$\xrightarrow{Evaluation}$Value Function$\xrightarrow{Greedy}$Policy then by generalized policy iteration.
To make is faster, value Iteration: use Bellman (optimality) operator to do iteration
But when $(\mathcal{P}, \mathcal{R})$ unknown, such as Go, gg
MC(episode)
First visit: sum from the first one to the last one: $R_{t+1} + \gamma R_{t+2} + ...\gamma^{T-t} R_T$
Every visit:
Both have convergence, but we need to run many times and when T is large the variance is large too.
Temporal Difference (TD):

$$v(S_t) \leftarrow v(S_t) + \alpha[R_{t+1} + \gamma v(S_{t+1}) - v(S_t)]$$

With probability 1, $\sum_t \alpha_t = \infty$ and $\sum_t \alpha_t^2 < \infty$ then it converges.
MC unbiased, high variance update utility at the end of episode; TD biased, low variance, every step update utility.
$TD(0) : v(S_t) \leftarrow v(S_t) + \alpha[R_{t+1} + \gamma v(S_{t+1}) - v(S_t)]$
$TD(n) : v(S_t) \leftarrow v(S_t) + \alpha[R_{t+1} + \sum_{i=1}^{n+1} \gamma^i v(S_{t+i}) - v(S_t)]$
Denote array μ_i as the index if $TD(i)$ now we want to update all the states at one time, that is:
$G_t^{(n)} := R_{t+1} + \gamma R_{t+2} ... + \gamma^n v(S_{t+n})$
$G_t^{\lambda} := (1 - \lambda) \sum_{n=0}^{\infty} \lambda^n G_t^{(n)}$