Main Idea

- Analyze DQN with NTK
- The main reason why DQN diverge(non-converge) is Over-generalization.
- Propose Pre-DQN to
- Pre-QN

Outline

- Main Ideas
- Analyzation Setup
- Building Intuition for Divergen with NTK
- Pre-QN
- Experiments

Your slide deck

Start writing!

HI

Loss Function

QMIX can be trained end-by-end, the loss function is defined as

$$L(\theta) = \sum_{i=1}^{b}[(y_i^{tot} - Q_{tot}(\tau, u, s; \theta))^2]$$

where b is the batch size of transitions sampled from the replay buffer, and $y_{tot} = r + \gamma\ max_{u'}\ Q_{tot}(\tau', u', s'; \theta^-)$, and θ^- are the parameters of a target network asin DQN