

Machine Learning Compendium

ML-KA Karlsruhe

August 29, 2018

Abstract

Looking up things take time. Let's just look into one single document

1 Reinforcement Learning

1.1 Bellman Equations

TODO backup diagrams
State Value Function

$$v_\pi(s) = \sum_{a \in A} \pi(s|a) Q_\pi(s, a) \quad (1)$$

Action Value Function

$$Q_\pi(s, a) = r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_\pi(s') \quad (2)$$

State Value Function recursive

$$v_\pi(s) = \sum_{a \in A} \pi(s|a) (r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_\pi(s')) \quad (3)$$

Action Value Function recursive

$$Q_\pi(s, a) = r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a \sum_{a' \in A} \pi(a'|s') Q_\pi(s', a') \quad (4)$$

Optimal State Value Function

$$v_*(s) = \max_a Q_*(s, a) \quad (5)$$

Optimal Action State Value Function

$$Q_*(s, a) = r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_*(s') \quad (6)$$

Optimal State Value Function recursive

$$v_*(s) = \max_a r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_*(s') \quad (7)$$

Optimal Action State Value Function recursive

$$Q_*(a, s) = r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a \max_{a'} Q_*(s', a') \quad (8)$$

1.2 Advantage Function

TODO

1.3 Policy, Policy Gradient

Policy: Distribution over actions given states

$$\pi_\theta(a|s) = P(a|s) \quad (9)$$

Policy Gradient

$$\nabla_\theta \pi_\theta(s|a) = \pi_\theta(s|a) \nabla_\theta \log \pi_\theta(s|a) \quad (10)$$

Note: this is valid for all probability distributions (the policy is a distribution over actions given states). The gradient term on the right hand side is called score function. The derivation basically uses the "log-trick".

2 Statistics

2.1 Probability

TODO: Probability (general + simple), CDF, Variance, L_p -Space for Random-Variables, Jensens-Inequality for Random Variables, Fisher-Information etc.

2.2 Distributions

TODO: All the usual suspects, including observed

2.3 Estimation

TODO: ML, Score-Function, biased/unbiased, Cramér–Rao bound, confidence-interval

2.4 Divergences

Conventions for this section: P and Q are probability measures over a set X , and P is absolutely continuous with respect to Q . S is a space of all probability distributions with common support.

2.4.1 Divergence

A divergence on S is a function $D : S \times S \rightarrow R$ satisfying

1. $D(p || q) \geq 0 \forall p, q \in S$,
2. $D(p || q) = 0 \Leftrightarrow p = q$

A divergence is a "sense" of distance between two probability distributions. It's not a metric, but a pre-metric.

[wikipedia](#)

2.4.2 f-Divergence

1. Generalization of whole family of divergences
2. For a convex function f such that $f(1) = 0$, the f-divergence of P from Q is defined as:

$$D_f(P || Q) \equiv \int_{\Omega} f\left(\frac{dP}{dQ}\right) dQ$$

3. [wikipedia](#)

2.4.3 KL-Divergence

1. The Kullback–Leibler divergence from Q to P is defined as
$$D_{KL}(P || Q) = \int_X \log \frac{dP}{dQ} dP = D_{t \log t}.$$
2. maximizing likelihood is equivalent to minimizing $D_{KL}(P(\cdot | \theta^*) || P(\cdot | \theta))$, where $P(\cdot | \theta^*)$ is the true distribution and $P(\cdot | \theta)$ is our estimate.

3. [wikipedia](#)

2.4.4 Jensen–Shannon divergence

The Jensen–Shannon divergence from Q to P is defined as

$$\text{JSD}(P || Q) = \frac{1}{2}D(P || M) + \frac{1}{2}D(Q || M)$$

where $M = \frac{1}{2}(P + Q)$

[wikipedia](#)

TODO: Wasserstein & Wasserstein Dual