Where is the uncertainty in Neural Networks?

Moritz Kniebel

Universität Tübingen

July 3, 2020

Why do we need uncertainty in Neural Networks?

Motivation

Problems

- NNs output point estimates
- Unknown uncertainties and overconfident
- Especially problematic in safety-critical applications (e.g. self-driving cars)

Motivation

Problems

- NNs output point estimates
- Unknown uncertainties and overconfident
- Especially problematic in safety-critical applications (e.g. self-driving cars)

Solutions

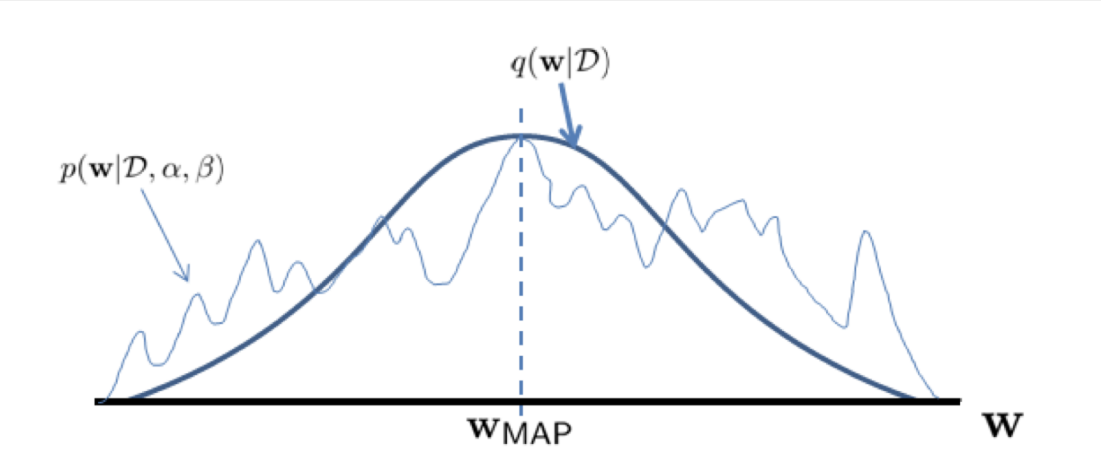
- Bayesian Neural Networks add a prior to the weights
- Posterior over weights can be formulated using Bayes theorem
- Posterior lets us make predictions about new data with a bound of confidence

Motivation

But...

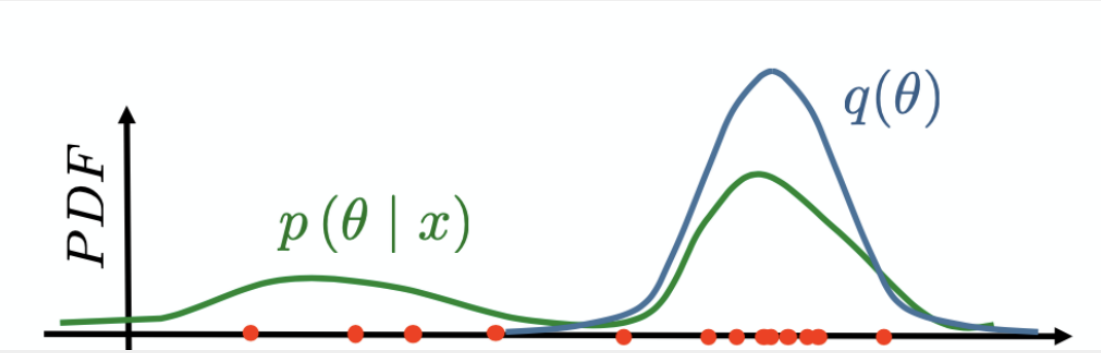
- Posterior over the weights becomes intractable
- Needs to be approximated

Laplace Approximation



- approx. true posterior by Gaussian centered at the mode of the weights W_{MAP}
- curvature is given by Hessian H w.r.t to the Loss L evaluated at W_{MAP}
- KFAC:
  - Hessian of every layer gets approximated by a Kronecker-product of two smaller matrices.

Variational Inference



- approx. true posterior $p(W|\mathcal{D})$ with parameterized variational distribution $q(W|\theta)$.
- objective: minimize the Kullback-Leibler divergence $\mathrm{KL}(q(W|\theta)||p(W|\mathcal{D}))$.
- tractable objective: maximize ELBO instead.

Research Goals

Research Goals

main objectives

- observe the differences in weight distributions to locate uncertainty
- locate the uncertainty in a single layer
- create visualizations of the uncertainty

extensions

- observe uncertainty during training

Possible implications

- training methods can focus on certain parts first
- unidentified parts could be pruned from a network
- Tracking uncertainty during training might give insights into convergence criteria (extension).

First goal

- Use network with simple architecture
- Apply Laplace approximation and Variational Inference to get uncertainty estimates.
- Find the location of the uncertainty
- create visualization tools to make findings more comprehensible

PROCEDURE

First goal

- Use network with simple architecture
- Apply Laplace approximation and Variational Inference to get uncertainty estimates.
- Find the location of the uncertainty
- create visualization tools to make findings more comprehensible

Second goal

- Use more complex network, such as VGG
- Transfer previous methods
- create visualization

possible extensions

- observe uncertainty during training
- measure the influence of the size of the weights to the resulting uncertainty
- add a third method (e.g. KFAC) to get uncertainty estimates

Results so far

TDB

- tbd
- tbd
- ...