ASSIGNMENT 2A: MAXIMUM LIKELIHOOD

Institute for Machine Learning

JOHANNES KEPLER
UNIVERSITY LINZ

Institute for
Machine Learning

Contact

Heads:
Thomas Adler,
Philipp Renz,
Andreas Radler

Institute for Machine Learning
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz

E-Mail: `{adler,renz,radler}@ml.jku.at`
Institute Homepage

Copyright statement:

Agenda

- Basics of Maximum Likelihood
- Properties of the Maximum Likelihood Estimator
- Further Literature:

\- Lecture notes (Hochreiter, 2014)

\- Mathematics for Machine Learning (Deisenroth et al., 2018)

Recap: Maximum Likelihood Estimator

- **Given:** n samples $\boldsymbol{x}_i$ from a distribution $p(\boldsymbol{x}; \boldsymbol{w})$, where $\boldsymbol{w} = \{w_1, w_2, \ldots, w_N\}$ represents the parameter(s) of the distribution.
- **Task:** find the parameter(s) $\boldsymbol{w}$ that were most likely used to produce this data.
- **Idea:** how likely was a given $\boldsymbol{w}$ used to produce the dataset?

$$\mathcal{L}(\{\boldsymbol{x}\}; \boldsymbol{w}) = \prod_{i=1}^{n} p(\boldsymbol{x}_i; \boldsymbol{w})$$

- **Solution:** find the $\boldsymbol{w}$ that maximizes $\mathcal{L}(\{\boldsymbol{x}\}; \boldsymbol{w})$ or $\ln \mathcal{L}(\{\boldsymbol{x}\}; \boldsymbol{w})$ (which is equivalent):

$$\hat{\boldsymbol{w}} = \operatorname{argmax}_{\boldsymbol{w}} \ln \mathcal{L}(\{\boldsymbol{x}\}; \boldsymbol{w})$$

Recap: Maximum Likelihood Estimator

Properties of the Maximum Likelihood Estimator (MLE) $\hat{\boldsymbol{w}}$:

- the MLE is invariant under parameter change
  - Model and data stay the same: Inference for two sets of parameters $\boldsymbol{w}$ and $\boldsymbol{\mu}$ should be the same, if $\boldsymbol{\mu}$ is a one-to-one transformation $\boldsymbol{\mu} = f(\boldsymbol{w})$, i.e. $\hat{\boldsymbol{\mu}} = f(\hat{\boldsymbol{w}})$

Recap: Maximum Likelihood Estimator

Properties of the Maximum Likelihood Estimator (MLE) $\hat{\boldsymbol{w}}$:

- the MLE is invariant under parameter change
- the MLE is asymptotically unbiased
  - Bias goes to 0 as sample size goes to infinity, i.e. $\lim_{n \to \infty} \mathrm{E}(\hat{\boldsymbol{w}}) = \boldsymbol{w}$

Recap: Maximum Likelihood Estimator

Properties of the Maximum Likelihood Estimator (MLE) $\hat{\boldsymbol{w}}$:

- the MLE is invariant under parameter change
- the MLE is asymptotically unbiased
- the MLE is asymptotically efficient, i.e. asymptotically optimal
  - Reaches CRLB as sample size goes to infinity, i.e. $\lim_{n\to\infty} \mathrm{Covar}(\hat{\boldsymbol{w}}) = \mathbf{I}_F^{-1}(\boldsymbol{w})$

Recap: Maximum Likelihood Estimator

Properties of the Maximum Likelihood Estimator (MLE) $\hat{\boldsymbol{w}}$:

- the MLE is invariant under parameter change
- the MLE is asymptotically unbiased
- the MLE is asymptotically efficient, i.e. asymptotically optimal
- the MLE is consistent for zero Cramer-Rao lower bound
  - If variance can go to 0 (i.e. zero CRLB) and MLE is asymptotically unbiased and asymptotically efficient, this means the estimated parameter will be the true parameter as the sample size goes to infinity, i.e. $\lim_{n \to \infty} \hat{\boldsymbol{w}} = \boldsymbol{w}$