

Expectation Maximization

Desmond Darma Putra

Overview

- Goal: find maximum likelihood solutions for models that having latent variables
- latent variables
 - variables that are not directly observed but are rather inferred from other variables that are observed (wikipedia)
- In K-means
 - E step: assigning the data point into a cluster
 - M step: re-compute the mean

EM Algorithm steps (1)

- Log likelihood function:

$$\ln p(X|\theta) = \ln \left\{ \sum_z p(X, Z | \theta) \right\}$$

- All observed data = X
- Latent variables = Z
- Set of all model parameters = θ
- Maximize $p(X|\theta)$ with respect to θ

EM Algorithm steps (2)

1. Choose an initial setting for the parameters θ^{old}
2. E-step: Evaluate $p(Z|X, \theta^{old})$.
3. M-step: Evaluate $\theta^{new} = \arg_{\theta} \max \varphi(\theta, \theta^{old})$
4. Where $\varphi(\theta, \theta^{old}) = \sum_z p(Z|X, \theta^{old}) \ln p(X, Z | \theta)$
5. Check the convergence in the log likelihood. If it is not satisfied then $\theta^{old} \leftarrow \theta^{new}$ and repeat step 2

Demo

- <http://jormungand.net/projects/misc/em/>

References

- Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- http://en.wikipedia.org/wiki/Expectation-maximization_algorithm