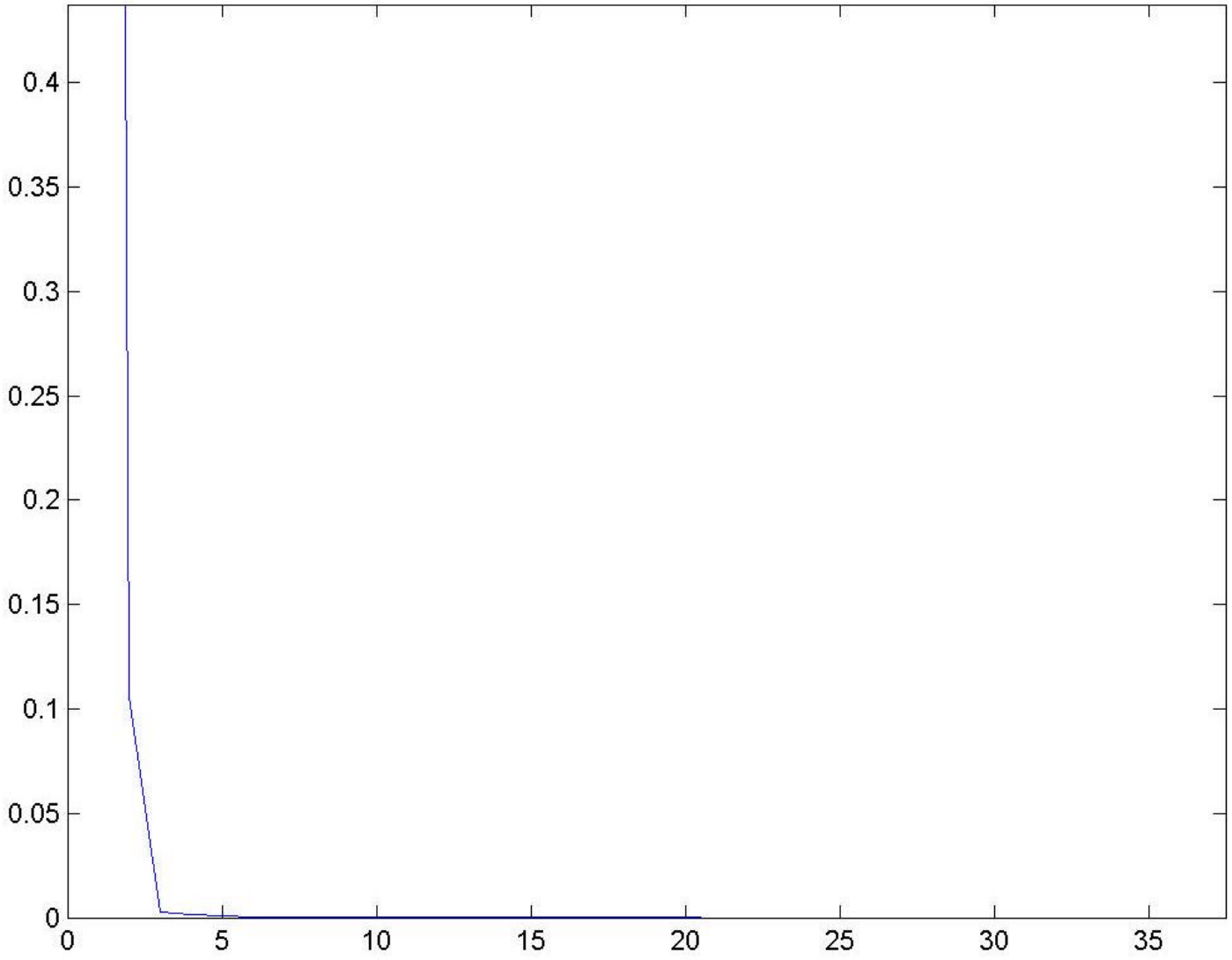
Datamining & Neural Networks

Session 3

PCA & ARD


0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
5
10
15
20
25
30
35

- Regularization on cost-function : $\beta E_D + \alpha E_W$

  e.g $E_W = 1/2 * \|w\|^2$

- three levels of inference

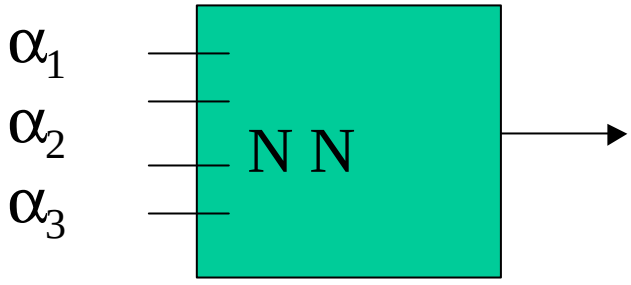
  1. Infer model parameters w ; α, β fixed
  2. Infer α, β ; given the model params

     $P(\alpha|D) \sim \mathbf{P(D}|a\mathbf{)} * P(\alpha)$
  3. comparison of different models

- **Evidence**-based framework of MacKay for estimation of hyperparameters : interpretation of
  - α as $1/\sigma^2_W$
  - β as $1/\sigma^2_D$
  - approx. of evidence factor yields update formula's for hyperparameters

ARD

- one hyperparam *per input* of your NN
- Goal: detection of most relevant inputs in the Bayesian framework



Costfunction = $\beta E_D + \alpha_1(\Sigma w_{i1}^2)+\alpha_2(\Sigma w_{i2}^2)+\ldots$

ARD in practice

1. Define NN model
2. Optimize weights
3. Optimize α, β , -- given weights
4. Cycle back to 2.

Exercise 2

- Demev1 (Evidence-based framework)
- Demard (ARD)
- Ionosphere dataset: 33 inputs, appr. 350 datapoints.
  - make classification NN (cfr. also session 2, ex. 3)
  - perform ARD, rank inputs
  - select subset of inputs based on α's
  - make new NN with this subset of inputs
  - compare performance with first NN