

#### Task:

Estimation of constant c, observed under the influence of noise  $e_k$ 

$$x_k = c + e_k$$

#### **Assumptions:**

- Noise is normally distributed with mean 0 and variance  $\sigma_e^2$
- Constant c is a random variable with zero mean and variance  $\sigma_c^2$

#### Given

- Prior knowledge of c ( $\mu_c$  , $\sigma_c$ ) and noise e (  $\mu_e$ ,  $\sigma_e$ )
- n observations of  $x_k \rightarrow$  training data

#### Goal:

"Good" estimation of c, taking into account the available knowledge sources.

#### Given

- Prior knowledge of  $c o p(c|\mu_c,\sigma_c^2)$
- Prior knowledge of  $e o p(e|\mu_e,\sigma_e^2)$

$$p(c|\mu_c=0,\sigma_c^2=5)$$

$$p(e|\mu_e = 0, \sigma_e^2 = 20)$$

#### Task:

Estimation of constant c, observed under the influence of noise  $e_k$ 

$$x_k = c + e_k$$

#### **Assumptions:**

- Noise is normally distributed with mean 0 and variance  $\sigma_e^2$
- Constant c is a random variable with zero mean and variance  $\sigma_c^2$

#### **Given**

- Prior knowledge of c ( $\mu_c$  , $\sigma_c$ ) and noise e ( $\mu_e$ ,  $\sigma_e$ )
- n observations of  $x_k$ --> training data

#### Goal:

"Good" estimation of  $c_{\rm r}$  taking into account the available knowledge sources.

#### Generation of training data

- Constant is assumed to have a fixed value:  $c=30\,$ 
  - --> prior knowledge is quite wrong!
- For different n ({1 2 4 8 16 32 64 128 256 512 1024}) Generate n training samples using  $x_k = c + e_k$
- For each training set with n samples:

Estimate and plot

- Log (likelihood):  $log(p(D|\mu))$
- Log (likelihood \* prior):  $log(p(D|\mu) \cdot p(\mu))$

Maximized for MAP parameter estimation

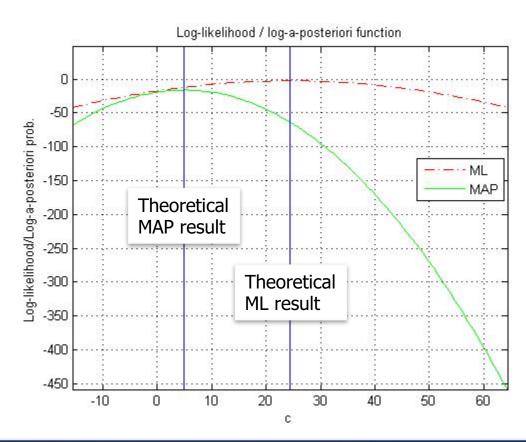
For each training set with n samples:

Estimate and plot log (likelihood), log (likelihood \* prior):

$$1. n = 1$$

e =

24.4132



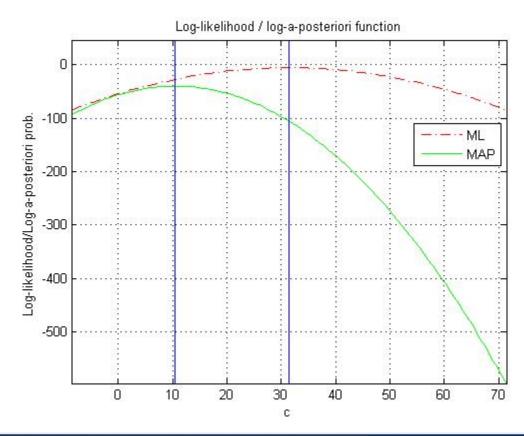
For each training set with n samples:

Estimate and plot log (likelihood), log (likelihood \* prior):

2. 
$$n = 2$$

e = 34.0240

28.8097



For each training set with n samples:

Estimate and plot log (likelihood), log (likelihood \* prior):

2. 
$$n = 4$$

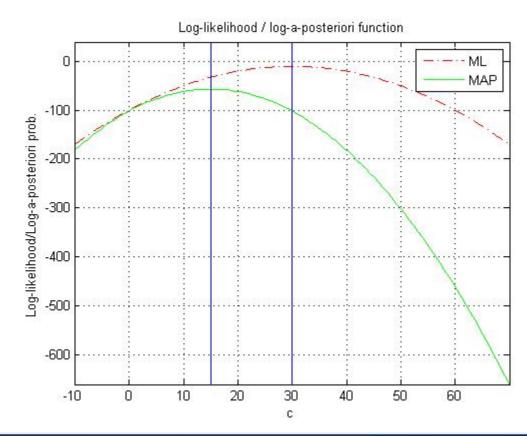
e =

34.5851

30.2318

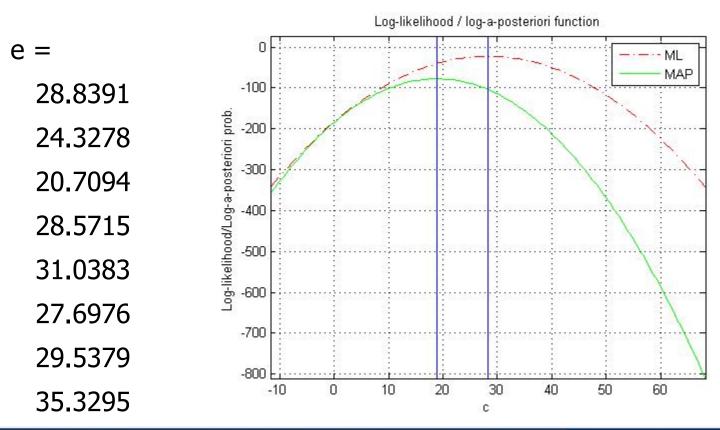
26.8810

28.0725



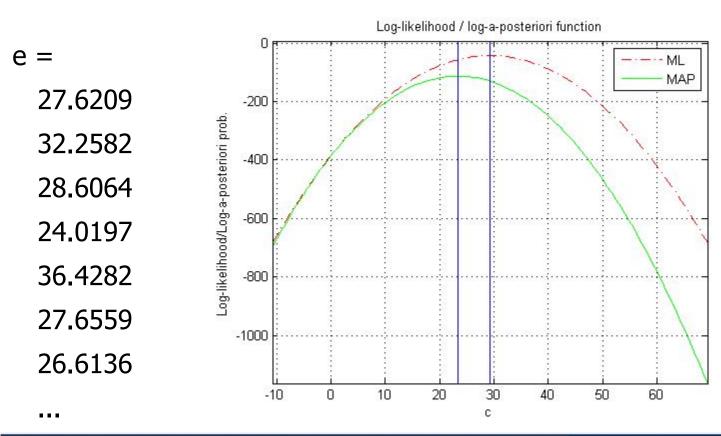
For each training set with n samples:

2. 
$$n = 8$$



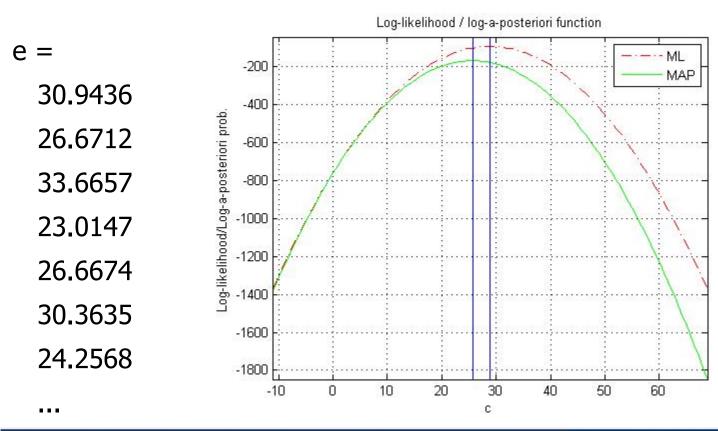
For each training set with n samples:

2. 
$$n = 16$$



For each training set with n samples:

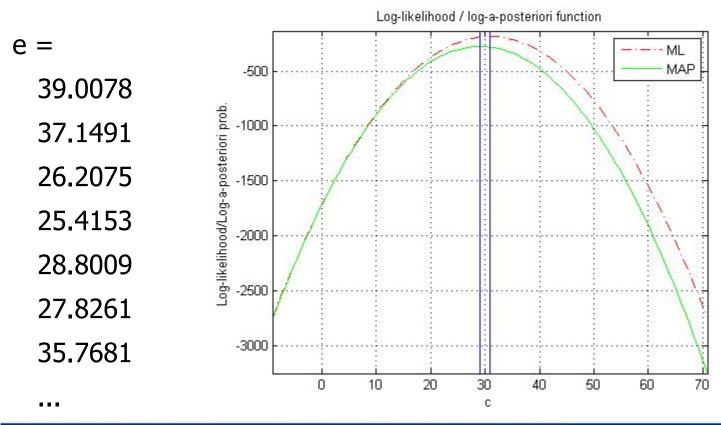
2. 
$$n = 32$$



For each training set with n samples:

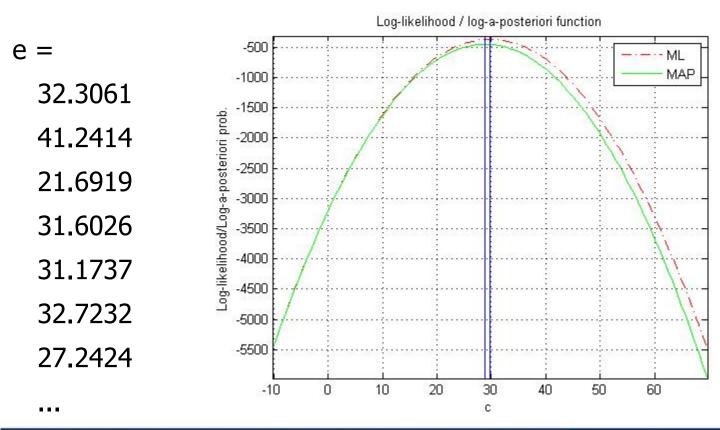
Estimate and plot log (likelihood), log (likelihood \* prior):

2. n = 64



For each training set with n samples:

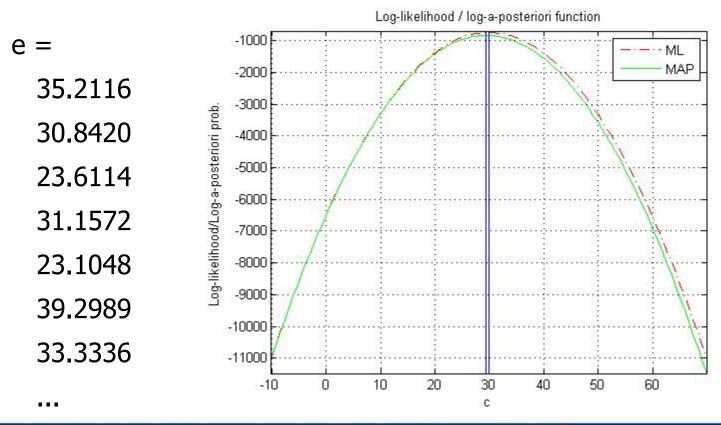
$$2. n = 128$$



For each training set with n samples:

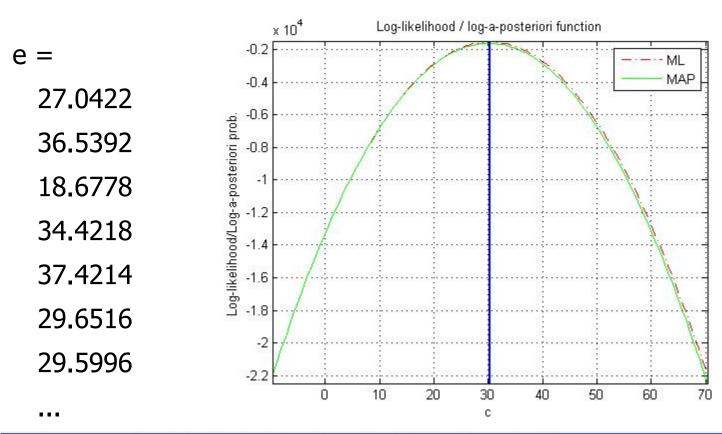
Estimate and plot log (likelihood), log (likelihood \* prior):

2. n = 256



For each training set with n samples:

2. 
$$n = 512$$



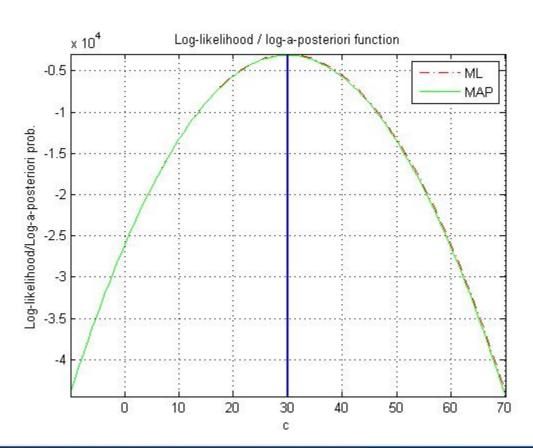
For each training set with n samples:

Estimate and plot log (likelihood), log (likelihood \* prior):

$$2. n = 1024$$

e =

28.3212
30.5653
33.1773
27.8407
21.0703
26.5033
26.0548



#### Result:

Both techniques have successfully estimated the "unknown" parameter c (which has been set to 30 for generation of the training data - see above).

