Adaptive Sampling for Model Selection

Nabil Hajj Chehade, Greg Pottie
EE, UCLA

Introduction

Model Uncertainty

- Sensor Networks are useful for learning about natural phenomena.
- Learning consists of extracting information from data to understand the underlying phenomenon.
- The purpose is to come up with a generalization of the phenomenon.
- Two basic components:
  1. Data collection or sampling,
  2. Modeling the data.
- Usually these two components are optimized separately.

Our idea

- Given some data, the model selection is optimized.
- Given a model, the sampling design is optimized.
- Our idea:
  Jointly optimize the two tasks.
- Interaction between sampling and model selection

Problem Formulation

Problem Formulation:

A set of 2 models:

$$H_1 : y_i = h_1(x_i, a) + e_i, \quad i = 1,\dots,n$$
$$H_2 : y_i = h_2(x_i, b) + e_i, \quad i = 1,\dots,n$$

Define:

Set of locations $Z = \{x_1 \ \cdots \ x_n\}$

Design $w = \begin{Bmatrix} x_1 & \dots & x_m \\ w_1 & \dots & w_m \end{Bmatrix}$

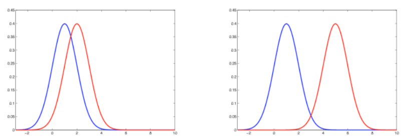
ML model selection, the probability of error is given by:

$$P = Q\left(\sqrt{\frac{\|h_1(x,a) - h_2(x,b)\|^2}{\sigma^2}}\right)$$

Assumptions:

The set contains *the correct model*.
Gaussian noise.

Idea: Find the locations that result in "the smallest" probability of error.



Mathematically:

$$\max_w \quad d^T w$$
$$\text{s.t.} \quad w \geq 0,\ 1^T w = 1$$

where:

$$d = \left(h_1(\mathbf{x}, \hat{a}) - h_2(\mathbf{x}, \hat{b})\right)^2$$
$$\hat{a} = \arg\min \left\|\mathbf{diag}(w)(y - h_1(\mathbf{x}, a))\right\|^2$$
$$\hat{b} = \arg\min \left\|\mathbf{diag}(w)(y - h_2(\mathbf{x}, b))\right\|^2$$

Interaction between sampling and estimation: we need *a* and *b* to find *w*, and we need *w* to find *a* and *b*. Hence the need of an iterative procedure to solve this problem.

Results

Algorithm

1.Given a design w_N, where N is the number of observations, find :

$$\hat{a}_N = \arg\min_a \sum_{i=1}^{N} (y_i - h_1(x_i, a))^2$$

$$\hat{b}_N = \arg\min_b \sum_{i=1}^{N} (y_i - h_2(x_i, b))^2$$

2. Add to the design a point x_{N+1} such that :

$$x_{N+1} = \arg\max_{z \in Z} \left(h_1(z;\hat{a}) - h_2(z;\hat{b})\right)^2 - c(z - z_j)^2$$

3.The $(N+1)th$ observation is taken at x_{N+1}

Update w : $w_{N+1} = (1-\alpha) * w_N + \alpha * \delta(x_{N+1})$

4.Go back to 1

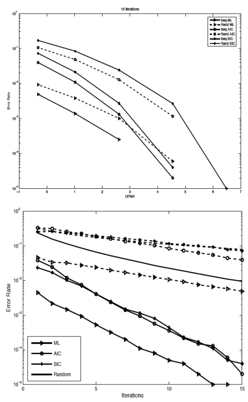
$$DPNR = \frac{\left\|h(x) - h_2(x;\hat{b})\right\|}{2\sigma}$$

Example:

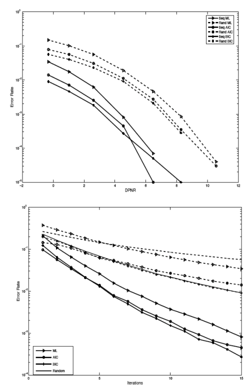
$$h_1(x_i;a) = a_0 + a_1 x_i \qquad i = 1,\dots,m$$
$$h_2(x_i;b) = b_0 + b_1 e^{x_i} + b_2 e^{-x_i} \qquad i = 1,\dots,m$$

Simulations:

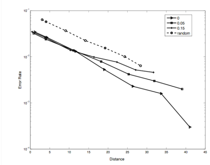
1. $h(x) = 4.5 - 0.6e^x - 1.3e^{-x}$



2. $h(x) = 2 + 5x$



3. Path planning



Future Work

- Convergence analysis of the algorithm
- Investigate the effect of the prior probabilities on the design.
- Extend the work presented to situations when the set of models considered does not include the correct model.
- Sampling robust to sensor faults (e.g. outliers)
- Spatio-temporal processes
- Sampling for non-parametric models.