Project instructions

Bayesian Computation

Guillaume Dehaene

Objectives

The objectives of the project are the following:

- Make sure you implement the methods presented in the class.
- Have you experiment with how to model a dataset.
- Look at real data.

It's possible to deviate from the standard project outline by discussing precisely with me what you want to do:

- I will send you a "contract" by email detailing the modified project instructions.
- You will have to respect this contract.

Steps of the project (1/3)

You need to do the following:

- Choose a dataset:
  - Analyzable with simple models:
    - Linear regression.
    - Linear classification.
    - Mixture of Gaussians.
  - Possible sources: Kaggle, classical datasets, SSC datasets (hyperlink).
  - Choose something that interests you if possible:
    - Finance.
    - Soccer.
    - League of legends / Dota.

Steps of the project (2/3)

- Implement a variety of models and approximation methods
  - At least three **qualitatively different** models (model = prior + likelihood).
    For example, for a model analyzable with a linear regression
    - Gaussian prior + Gaussian noise (very basic model).
    - Sparsity-inducing prior + Gaussian noise (parameters are probably sparse).
    - Gaussian prior + student noise (possible outliers in the data).
  - At least three **qualitatively different** approximation methods including Metropolis-Hastings. E.g.
    - Metropolis-Hastings.
    - Importance sampling.
    - Gaussian Variational approximation.
- **Make your code readable ! You will have to submit it !**

Steps of the project (3/3)

- Comparison of the models and methods:
  - Which model generalizes bests?
  - How much time does each approximation method require?
  - Which method do you feel was the best for your problem and why?
- Presentation

Presentation

You should answer the following questions:

- What important features does you data have? (1-2 slides).
- What models did you use and what do they bring you? (from 1 slide to 1 slide per model)
- What approximations did you use? How do they work (briefly)? (1 slide per approximation)
- Result of the comparisons.

- Which method is the best on your problem?
  What features of your problem matter here?

General advice

- Come to the exercise sessions so that you can implement the methods here.
- **Seek a lot of feedback about your project**.
- Do not start working on the project in week 10.
- Work in groups but do not copy your code.
- **You should implement all approximation methods yourself. I will check your code for this.**
- **COME TO ALL THE EXERCISE SESSIONS**.