Due Date: 17:00, Fri, March 13, 2015
TA: Yi Shen (yishen@sep.stanford.edu)

Lab 7 - Hyperbolic model fitting

Your name[1]

ABSTRACT

You've almost made it! This will be the final lab and is designed to group together almost all of the theory you've learnt over the last two months into a practical problem using real data. We'll make use of helical coordinates, leaky integration, polynomial division, ℓ_2 based minimisation, ℓ_1 based minimisation and the hyperbolic penalty function to create a bathymetry map of the bottom of the Sea of Galilee. The main goals herein are to increase awareness with how we can create a strategy to deal with data inadequacies and the performance of different norms.

INTRODUCTION

Hyperbolic / hybrid model fitting

Chapter 7 in the book presents an outline of the hyperbolic model fitting problem. In a least-squares fitting problem the penalty function is a hyperbola, in an ℓ_1 ftting problem the penalty function looks like an absolute valued function. Thus we can design a hybrid penalty function that transitions from ℓ_2 to ℓ_1 at a certain threshold r_i to be determined.

The Galilee dataset

This dataset is a collection of sonar readings taken from various surveys over the Sea of Galilee, in total we have 132,044 measurements. This data-set is small enough that we can run our basic codes without much waiting time, but the underlying geophysics complex enough that we can learn a lot about dealing with real data. The three mains problems are: the data is irregular, there are various types of noise bursts, and there is

[1]**e-mail:** youremail@stanford.edu

systematic error in the form of drift. These three attributes are found in the majority of data-sets you will have to deal with in geophysical problems.

Our goal is take these data and create a bathymetry map of the bottom of the lake, but we need to think about how to deal with these various problems. We want to avoid manually searching for errant and/or inconsistent values because for a large dataset this is unfeasible - rather we want to design our inversion to help overcome the problems in the data.

YOUR ASSIGNMENT

This assignment will not be as heavy on the coding as the previous few, but will rather get you to focus on concepts and results. You are provided with a set of codes that you should read and become familiar with, especially `Galilee.f90`, you will be using this to create your results. You have been provided with some plotting rules to create your figures.

Questions

Norms

1. What is the difference between the ℓ_1 norm and the ℓ_2 norm?

YOUR ANSWER:

2. How do large residuals affect the ℓ_2 norm?

YOUR ANSWER:

3. How do large residuals affect the ℓ_1 norm?

YOUR ANSWER:

4. If solver issues were not a problem, which norm would you prefer to use for the interval velocity estimation problem? Why?

YOUR ANSWER:

5. When using the hybrid norm, how can we decide where to place our threshold? Is there a quantitative way of determining this?

YOUR ANSWER:

Hybrid norm and binning

6. Look at the codes provided. How do we input the L1/L2 threshold parameter? Which code(s) use this parameter and how?

YOUR ANSWER:

7. Open the Makefile and run the rule that creates the data and the target `l2_noint.H` and `Fig/initial_data.v`. Write a rule that follows the target `hyper_noint.H` that will run the inversion using the hybrid solver, you'll have to provide the code with the threshold parameter, use around 0.01 for now. Use the rule `make Fig/hybrid_vs_l2.v` to produce a figure that compares the two. What are the key differences between the images? Include the figure here. Elaborate. What features do the annotations point to?

YOUR ANSWER.

8. Also create and include the figure `xsections.v`. What does this figure display? Does this view provide any further insight to your answer for the previous question?

YOUR ANSWER.

9. We used a threshold parameter of 0.01 previously. Does this seem reasonable? Experiment with some other parameters and use a similar makerule to `hybrid_vs_l2.v` to include them here. What happens for very large thresholds? For very small thresholds? Should this solver work correctly close to the ℓ_2 limit? Close to the ℓ_1 limit?

YOUR ANSWER.

10. You were already provided with some binning parameters. Try a coarser grid and a finer grid and show the differences that this makes to this L2 and hybrid result. Explain where these come from and how they affect the inversion. Include a figure comparing a coarser grid with a finer grid.

YOUR ANSWER.

Removing the acquisition footprint

11. The figures already generated should show you that there are many acquisition footprints remaining. Complete the rule for creating `hyper-dds.H` and `Fig/dds.v`. What does this inversion do? Why might it help remove some of the track artifacts?

YOUR ANSWER.

12. Include the figure you created. Has it done a good job of removing the acquisition? What problems still remain? Is there anything else we could do to this result to improve it? If not, then why?

YOUR ANSWER.

Modeling the drift

13. We will now attempt to additionally quantify the drift. Write out the fitting goals we are now trying to solve in the usual way, let $\mathbf{L}$ be a drift modeling operator and τ a balancing parameter. Also write the new misfit function we are aiming to solve. Why is estimating the drift necessary?

YOUR ANSWER.

14. We now have an additional variable to be estimated - is this possible? What problems could this cause?

YOUR ANSWER.

15. What operator could **L** be? Speculate on 2-3.

YOUR ANSWER.

16. We decide to try leaky integration, complete the rule in the makefile to do this and decide on a few values of ρ for comparison. You should be able to understand the make rules beginning with `wo_tracks` and `resids` and use them to create 3 image comparisons with different values of ρ. Include your best one here, what was the value of ρ?. What aspects are better than the simple hybrid approach? What are worse?

YOUR ANSWER.

17. The parameter τ balances the modeling of the drift with the rest of the inversion. Experiment with a few values - which gives the best result? If you can improve on your previous result, show the figure here.

YOUR ANSWER.

18. Is there any way to decouple the drift estimation from the rest of the data fitting? If not, why not?

DONE

When you are all finished modifying the source files, `Makefile` and latex file, make sure that will paper will compile correctly from a cleaned directory using the default targets only. In other words, the following sequence of commands should produce your a PDF version of your paper without any problems: `make burn`; `scons -c`; `make`; `scons`.

Once you make sure everything is working properly, clean up your directory by typing `make burn`; `scons -c`. Then, compress the lab directory using the command `tar cvzf Lab7.gzip Lab7` and submit the compressed file to your TA.