

Project Proposal: Graphical models for precipitation

Zachary DeStefano, 15247592

Homer Strong, 94767004

Disi Ji, ????????

CS 274B: Spring 2016

Introduction

Precipitation plays a major role in many natural and societal systems, but direct observations of precipitation are not available for all areas. Remote sensing platforms such as satellites provide estimates of temperature via measurements of electromagnetic radiation. Temperature and precipitation are closely related due to the physical mechanisms which govern precipitation. Furthermore there is strong spatial structure in both temperature and precipitation, so the ability to describe spatial dependence is highly desirable for models of precipitation. We consider modelling precipitation given temperature on a regular grid of pixels, where the values of each pixel corresponds to average amounts over a fixed time window. Graphical models provide a natural approach for taking into account the spatial dependencies inherent in patterns of precipitation.

Precipitation dataset

We will use PERSIAN-CCS data set which consists of temperature data as well as features related to a particular cloud patch in which a pixel is part. Each pixel, which represents a small geographic area, has 1 temperature feature and 12 features related to its cloud patch. There is also a target y value for that pixel obtained via radar data. The region of interest is the western United States and the temporal resolution under consideration is the finest available, 30 minutes. The following web link provides a more detailed description of the data set.

http://chrs.web.uci.edu/research/satellite_precipitation/activities01.html

Graphical models

In order to predict precipitation, we will use a Markov Model where each x_i is conditioned on each y_i . Each y_i is conditioned on its neighbors. In this way, neighboring pixels will be likely to have similar values. We will try a few different exponential family models for the relationship between x_i and y_i . For the relationship between neighboring y_i values, we will also use exponential family models and potentially Conditional Autoregressive models. We will try both a single parameter set for all the pixels as well as separate parameter sets for each pixel. We will evaluate our models based on their error when applied over a test data set.

questions: how to model continuous relationship? how to perform model selection? model occurrence?

Expected outcomes

The goal for this project is to evaluate several graphical models for precipitation given temperature over a grid of pixels covering the western United States. We intend to focus on the spatial structure, so no temporality will be included in these models. The selected model should accurately predict precipitation using temperature data.