Near Optimal sensor placement for precipitation measurement

Kei Usui
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213
kusui@andrew.cmu.edu

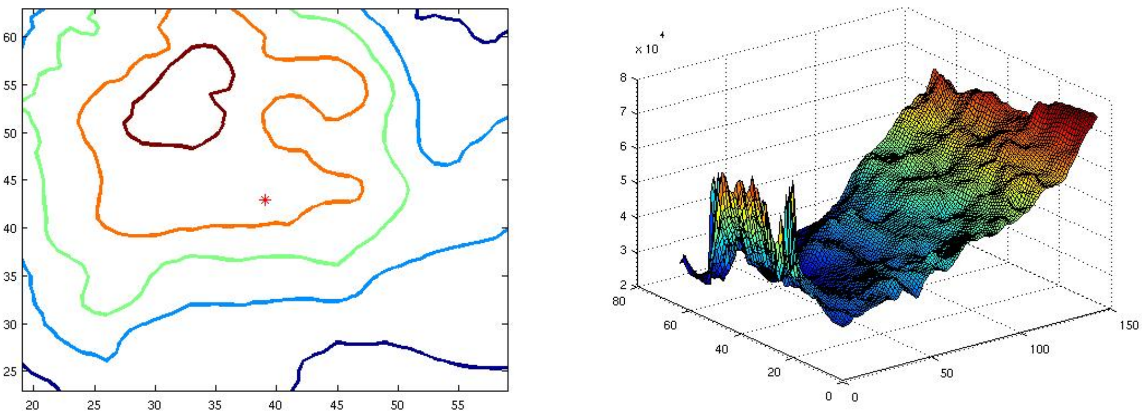
Abstract

Selecting the best locations for sensors is a very important task when monitoring spatial phenomena. Gaussian processes are often employed to avoid strong assumptions such as fixed sensing radii when modeling the underlying phenomena. In this paper, a mutual information criterion is used to choose sensor locations which most reduces uncertainty about unsensed locations. An approximation to this NP-complete problem is introduced. Finally, sensor placements selected by this algorithm is compared to placements using formerly popular entropy criterion.

1 Introduction

Selecting the best locations for sensors is a very important task when monitoring spatial phenomena. One approach is to assume that sensors have a fixed sensing radius, however as shown in Fig. 1(a), in practice this assumption is often too strong. Sensing area is largely affected by the surrounding environment and is usually not characterized by a regular disk. Alternative approach to this is learning a Gaussian Process (GP) model for the phenomena.

Typical sensor placement techniques place sensors greedily at the highest entropy location of the GP where uncertainty about the phenomena is largest. Problem with this approach is that entropy is an indirect criterion in the sense that it does not consider the prediction quality of the areas around the selected placements. Mutual information criterion directly measures the effect of sensor placements on the posterior uncertainty of the GP and seeks to find sensor placements that are most informative about unsensed locations. In this paper, an approximation algorithm to this NP-complete problem is applied to find near-optimal sensor locations to measure precipitation in Pennsylvania (Fig. 1(b)).



a) non-radial covariance distribution b) average annual precipitation in Pennsylvania

Figure 1.a) covariance of precipitation at star (40°21N,76°29W) and region around

b) Average annual precipitation in Pennsylvania (milliliters) in the years 1990-1999

2 Method

2.1 Gaussian Process

A Gaussian Process (GP) is employed to model the precipitation estimates at points where no sensors are placed. Even though linear regression often gives excellent predictions, there usually is no concept of uncertainty while GP is a non-parametric generalization of linear regression that allows to represent uncertainty. Non-isotropic GP marginal distribution is given as

$$p(\mathbf{y}) = N(\mathbf{y} \mid 0, \mathbf{K})$$

where $\mathbf{K}$ is chosen to express the property that, for points $\mathbf{x_n}$ and $\mathbf{x_m}$ that are similar, the corresponding values $\mathbf{y_n}$ and $\mathbf{y_m}$ will be more strongly correlated than for dissimilar points. Details of $\mathbf{K}$ used in this paper is given in later section. Marginal distribution for observed value $\mathbf{t}$ is given in very similar form

$$\mathbf{t} = \mathbf{y} + \text{noise}$$

$$p(\mathbf{t}) = N(\mathbf{t} \mid 0, \Sigma)$$

$$\Sigma = K + B^{-1}$$

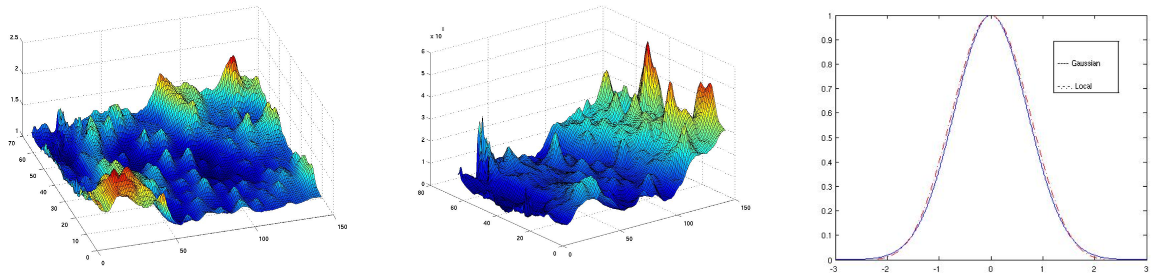
where B represents the precision of the noise. Conditional Gauss distribution given set of sensor measurements X_A where A is the set of sensor locations, is then given by linear-Gaussian model

$$\mu_{\mathbf{y|A}} = \mu_{\mathbf{y}} + \Sigma_{\mathbf{yA}}\Sigma_{\mathbf{AA}}^{-1}(\mathbf{x_A} - \mu_{\mathbf{A}}) \quad (1)$$

$$\Sigma_{\mathbf{y|A}} = \Sigma_{\mathbf{yy}} - \Sigma_{\mathbf{yA}}\Sigma_{\mathbf{AA}}^{-1}\Sigma_{\mathbf{Ay}} \quad (2)$$

2.2 Optimizing sensor placements

An intuitive criterion for finding sensor placements is to choose A which maximizes entropy H(A) since the set of sensors which covers the space where other sensors are uncertain about should cover the space well.



a) mutual information b) uncertainty c) local kernel

Figure 2: a) Mutual information, b)Uncertainty before placing any sensors. c) Gaussian distribution and resembling local kernel

Conditional entropy H(Y | A) represents the uncertainty of precipitation at Y given set of sensors A and is given by

$$H(Y \mid A) = (\log \Sigma_{Y|A} + \log(2\pi)+1) / 2 \quad (3)$$

where $\Sigma_{Y|A}$ is given by (2). However, this criterion suffers from the fact that it only considers the entropy of the locations where sensors are actually placed but not the entire space of interest. An alternative criterion used in this paper takes the subset of sensor locations that most significantly reduces the uncertainty about the estimates in the rest of the space. Specifically, quantity H(V\A) – H(V\A | A) is maximized where "V\A" denotes a set consisting of elements of V which are not in A. H(V\A) represents the uncertainty of unsensed locations and H(V\A | A) represents the uncertainty remaining in those locations given sensors A. Thus, mutual information between A and the rest of the space, H(V\A) – H(V\A | A), can be seen as how much uncertainty is removed in the unsensed locations by placing set of sensors A. The term -H(V\A | A) forces the algorithm to select sensor locations away from the boundaries of unsensed locations. Maximizing either entropy or mutual information is known to be NP-complete ([3] Ko et al. (1995)).

2.3 Approximation Algorithm

The algorithm used is greedy, repeatedly adding sensors which maximize increase in mutual information. For further computational feasibility, it is assumed that precipitations of two points far away are independent. To accomplish this, local kernels proposed by [4] Storkey (1999) is employed as shown in Fig. 2(c).

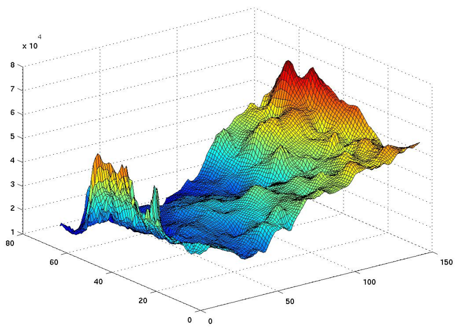
$$K(x,y) = ((2\pi-\Delta)(1+(\cos(\Delta)/2)+3/2\sin(\Delta)) / 3\pi \quad (4)$$

for $\Delta<2\pi$ and 0 otherwise, where Δ is the euclidean distance between x and y times some positive constant. This local kernel resembles Gaussian distribution for $\Delta<2\pi$ and total probability in the range $[-2\pi,2\pi]$ adds up to one.

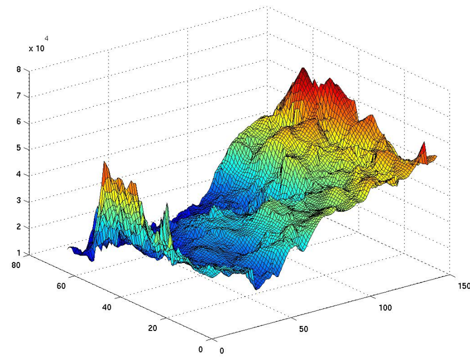
Given number of sensors k, space of interest V, the algorithm returns sensor selection A which is a subset of V as follows. Initialization involves computation of mutual information at every point in V as shown in Fig. 2(a). Then repeat for k times, add the point with highest mutual information into A and update mutual information of the points within the local kernel of newly added sensor location as shown in Algorithm. In the Algorithm pseudocode, notation "V\Y" is used to mean rest of space within local kernel instead of rest of entire space for the reason discussed earlier. Also, LK(Y) stands for set of locations within the local kernel of Y where (4) is greater than zero.

```
A ← [];
foreach Y ∈ V do
MI_Y = H(V\Y) – H(V\Y | Y)  //initialization
for j = 1 to k do
        Y* = arg max_Y MI_Y  //greedy placement of next sensor
        A ← A ∪ Y*
        foreach Y ∈ LK(Y*) do
                MI_Y = H(V\Y | A) – H(V\Y | A∪Y)    // update mutual information given updated A
return A
end
```

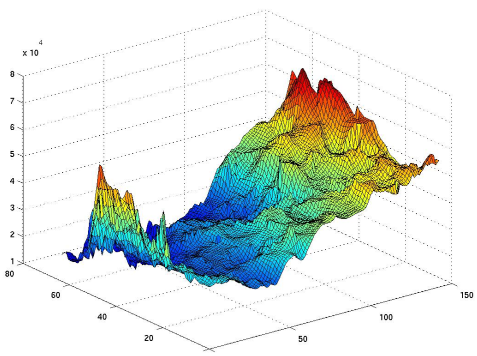
Algorithm: Pseudocode for greedy approximation algorithm



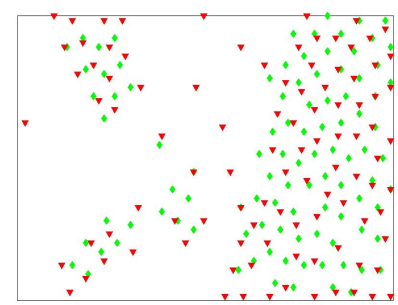
a) Mutual information



b) Entropy



c) Actual precipitation

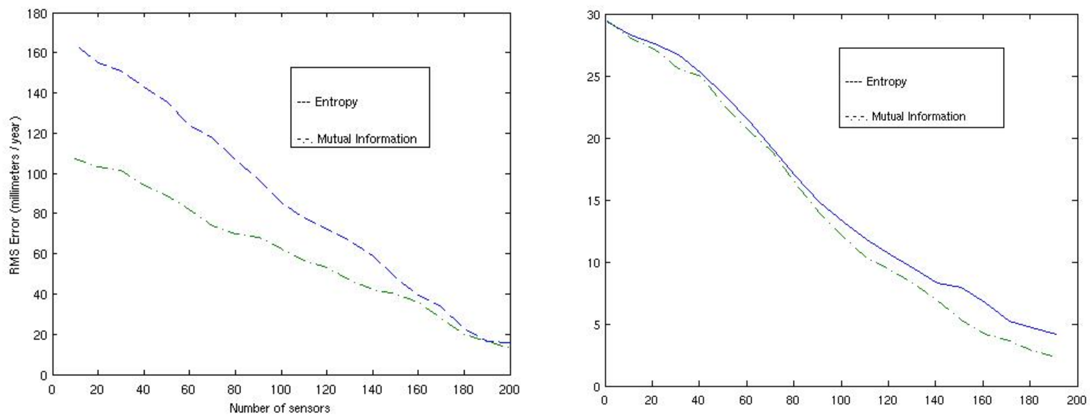


d) Sensor locations

Figure 3: a,b) predicted precipitation using 200 sensors for year 2001
c) actual precipitation for year 2001
d) Sensor locations (k=100), squares are mutual information and triangles are entropy

3 Experiments

Annual precipitation data collected during the years 1990-1999 in 9152 regions of equal area approximately 4 km apart in the state of Pennsylvania ([1] Widmann and Bretherton) is used to test the approach. Data was preprocessed in order to ensure that it can be reasonably modeled using GP. Log-transformation was applied first to precipitation values and then annual mean was subtracted to for more accurate covariance information.



a) Average RMS b) Average Error rate (%)

Figure 4: a) Average root-mean-squares error for the years 2001-2005
b) Average error rate for the years 2001-2005

Given observed values at sensed locations, predicted values at unsensed locations are computed using (1). When there is no sensor within the local kernel of unsensed locations, simply mean value from the training data set μ_y was used.

It can be observed from Fig. 3(d) that entropy criteria tends to put sensors along the boundary much more than mutual information criteria as predicted in earlier section. Fig. 4(a) shows average root-mean-squares error for the years 2001-2005. Fig. 4(b) shows error rate given by ratio of absolute difference of predicted and actual precipitation, and actual precipitation. In either measurement, mutual information outperforms entropy as a selection criteria by not placing sensors near boundary and wasting their sensing radii.

4 Conclusions

In this paper, I (i) showed how mutual information and entropy perform as a criterion for sensor placement in Gaussian Process, (ii) evaluated performance of approximation algorithm to optimizing either criterion. It was shown how mutual information as a criterion for sensor placement to monitor spatial phenomena outcomes entropy.

5 Reference

[1]
Widmann, M. and C. S. Bretherton, 2000: Validation of Mesoscale Precipitation in the NCEP Reanalysis Using a New Gridcell Dataset for the Northwestern United States.

[2]
Carlos Guestrin, Andreas Krause and Ajit Singh, June 2005: Near-Optimal Sensor Placements in Gaussian Processes , CMU- CS-05 -120, CMU- CALD- 05-102

[3]
Ko, C., Lee, J., & Querance, M. 1995: An Exact Algorithm for Maximum Entropy Sampling, Ops Research, 43, 684-691.

[4]
Storkey, A.J. 1999: Truncated Covariance Matrices and Toeplitz Methods in Gaussian Processes, ICANN99.
[5]
Bishop, C.M. 2006: Pattern Recognition and Machine Learning