Seminar Presentation:

Active Data Selection For Gaussian Process Regression

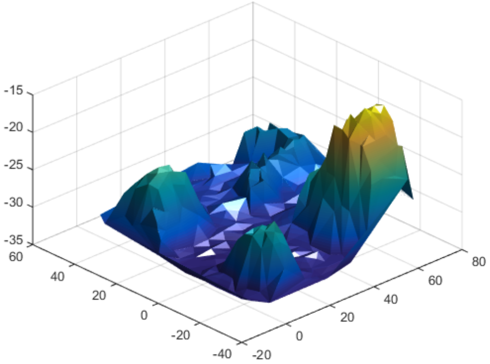
Nate Schambach
Group 10
Mentors: Prof. Kobilarov, Prof. Taylor, Preetham Chalasani

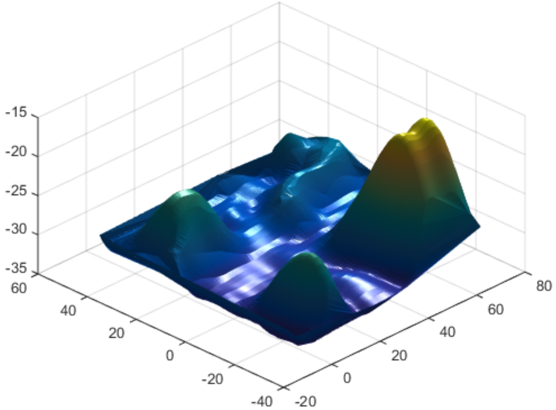
LABORATORY FOR Computational Sensing + Robotics

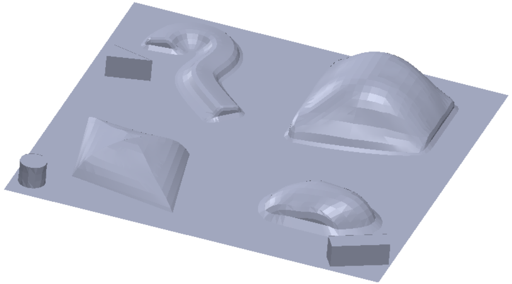
ERC | CISST

Optimized Tissue Reconstruction

Geometry reconstruction of tissue using minimal number of points.







Presentation Outline

Background
Paper and Motivation
Active Data Selection
Test Point Rejection
Thoughts

Some GPR Background

- Gaussian Process
  - A collection of random variables that have joint gaussian distributions

$$P(\boldsymbol{t}|\boldsymbol{C},\boldsymbol{x_n}) = \frac{\mathbf{1}}{Z}\exp\left(-\frac{1}{2}(\boldsymbol{t}-\mu)^T\boldsymbol{C}^{-1}(\boldsymbol{t}-\mu)\right)$$

  - Prediction:

$$\hat{y}(\tilde{x}) = \boldsymbol{k}(\tilde{x})\boldsymbol{C}_N^{-1}\boldsymbol{t}$$
$$\sigma_{\hat{y}}^2(\tilde{x}) = C(\tilde{x},\tilde{x}) - \boldsymbol{k}(\tilde{x})\boldsymbol{C}_N^{-1}\boldsymbol{k}(\tilde{x})$$

Paper Selection

- Seo, S., Wallat, M., Graepel, T., Obermayer, K., *Gaussian Process Regression: Active Data Selection and Test Point Rejection.* Department of Computer Science, Technical University of Berlin, 2000.

- Project goal: Accurate And Efficient Tissue Reconstruction, paper helps us choose the fewest points to palpate.

The Problem:

- Not all points are created equal.
  - Which point will give us the most information
  - Should some points be rejected?

Key Results:

Minimization of Average Variance drastically accelerates learning

Throwing out points also accelerates learning but less so without an accurate model

Active Learning McKay (ALM)

Select X points to predict values for.

↓

Calculate their expected value and variances.

↓

Choose point with maximum variance to sample next.

Active Learning Cohn (ALC)

Minimization of Generalization Error: $E_{MSE} = \sigma_{\hat{y}}^2 + E_x[(E_\tau[\hat{y}(x)] - y(x))^2]$

Compute how the overall variance would change for X points:

$$\mathbf{C}_{N+1} = \begin{bmatrix} \mathbf{C}_N & \boldsymbol{m} \\ \boldsymbol{m}^T & C(\tilde{x}, \tilde{x}) \end{bmatrix} \mathbf{C}_{N+1}^{-1} = \begin{bmatrix} \left[\mathrm{C}_N^{-1} + \frac{1}{u} \boldsymbol{g}\boldsymbol{g}^T \right] & \boldsymbol{g} \\ \boldsymbol{g}^T & u \end{bmatrix}$$

$$\boldsymbol{m} = [C(x_1, \tilde{x}) \ldots C(x_N, \tilde{x})] \in \mathbb{R}^N$$

$$\boldsymbol{g} = -u\boldsymbol{C}_N^{-1}\boldsymbol{m}, \qquad \boldsymbol{u} = \left(C(x_N, \tilde{x}) - \boldsymbol{m}^T \boldsymbol{C}_N^{-1} \boldsymbol{m} \right)^{-1}$$

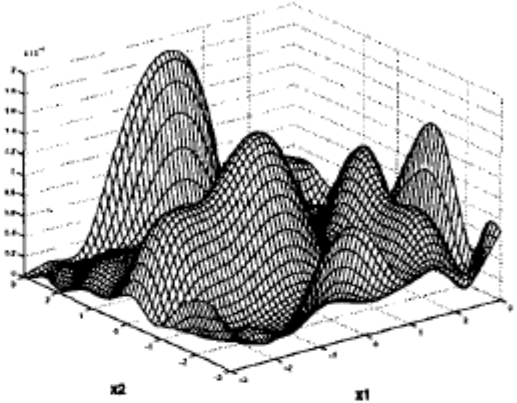


Fig. 1.b. Seo

Choose the point with the largest change in the overall variance.

$$\Delta\sigma_{\hat{y}(\xi)}^2(\tilde{x}) = \sigma_{\hat{y}(\xi)}^2 - \sigma_{\hat{y}(\xi)}^2(\tilde{x}) = \frac{\left(\boldsymbol{k}_N \boldsymbol{C}_N^{-1} \boldsymbol{m} - C(\tilde{x}, \xi) \right)^2}{(C(\tilde{x}, \tilde{x}) - \boldsymbol{m}^T \boldsymbol{C}_N^{-1} \boldsymbol{m})}$$

Active Learning Cohn (ALC)

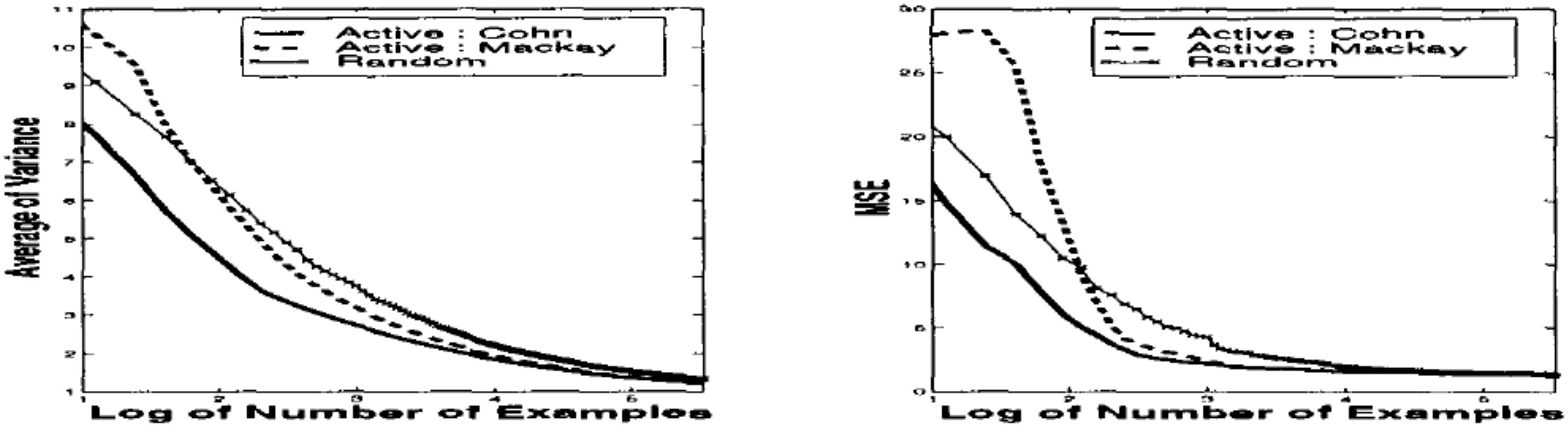


Fig 4 Seo.

Test Point Rejection

Compare your predictions at the values you have tested and remove those which are causing a poor fit.

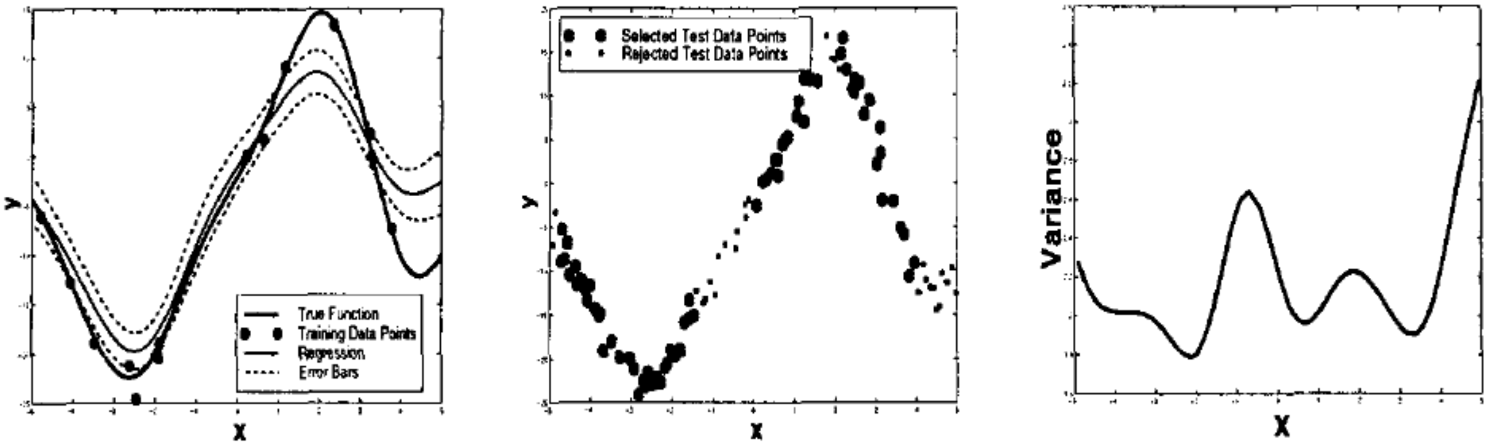


Fig.2 Seo

Test Point Rejection

Works much better when the model you are using is closer to the true model.

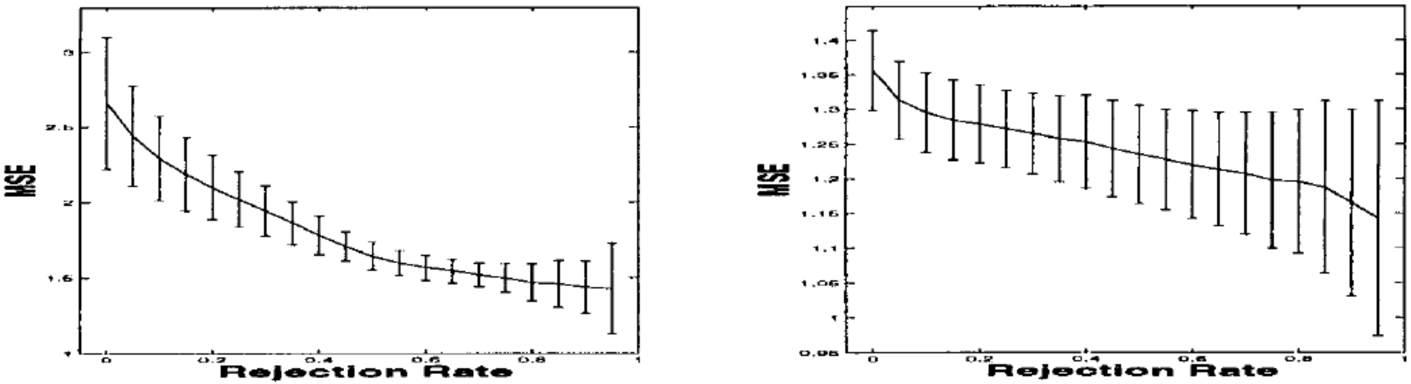


Fig. 5 Seo

Final Thoughts

- Test Point Rejection: a welcome but unexpected addition
- ALC is an effective alternative to ALM
- Better evaluation of what is the “best” improvement; Are there other methods than ALC for a different “best” improvement?

Questions?