Sparse Cholesky Factorization by Greedy Conditional Selection
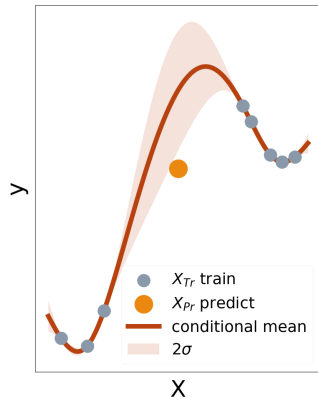
Stephen Huan   Florian Schäfer

Georgia Institute of Technology

February 19, 2022

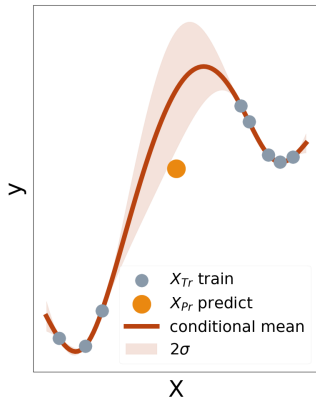# The Problem: Gaussian Process Regression

Measurements $\boldsymbol{y}_{\mathsf{Tr}}$ at N points X_{Tr}

The Problem: Gaussian Process Regression

Measurements $\boldsymbol{y}_{\mathsf{Tr}}$ at N points X_{Tr}

Estimate unseen data $\boldsymbol{y}_{\mathsf{Pr}}$ at X_{Pr}

The Problem: Gaussian Process Regression

Measurements $\boldsymbol{y}_{\mathsf{Tr}}$ at N points X_{Tr}

Estimate unseen data $\boldsymbol{y}_{\mathsf{Pr}}$ at X_{Pr}

Model as Gaussian process
$\rightarrow$ condition on $\boldsymbol{y}_{\mathsf{Tr}}$
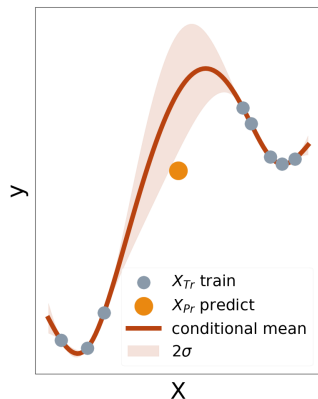
The Problem: Gaussian Process Regression

Measurements $\boldsymbol{y}_{\mathsf{Tr}}$ at N points X_{Tr}

Estimate unseen data $\boldsymbol{y}_{\mathsf{Pr}}$ at X_{Pr}

Model as Gaussian process
$\rightarrow$ condition on $\boldsymbol{y}_{\mathsf{Tr}}$

Computational cost scales as N^3
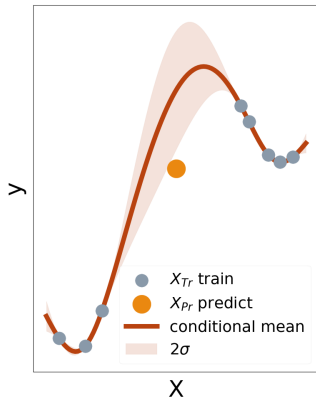
The Problem: Gaussian Process Regression

Measurements $\boldsymbol{y}_{\mathsf{Tr}}$ at N points X_{Tr}

Estimate unseen data $\boldsymbol{y}_{\mathsf{Pr}}$ at X_{Pr}

Model as Gaussian process
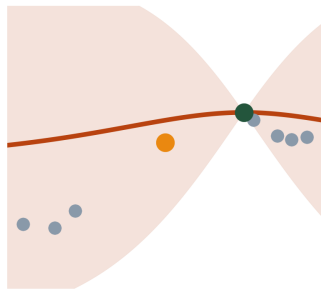$\rightarrow$ condition on $\boldsymbol{y}_{\mathsf{Tr}}$

Computational cost scales as N^3

Choose k most informative points!
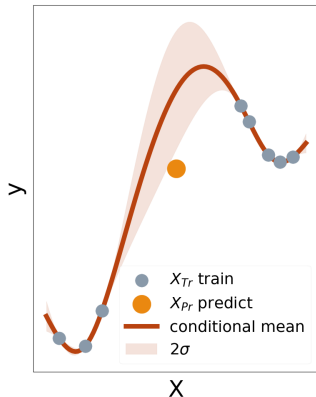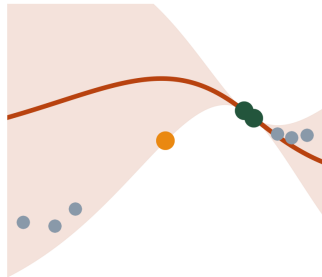
Conditional k-th Nearest Neighbors

Naive: select k closest points

Conditional k-th Nearest Neighbors

Naive: select k closest points

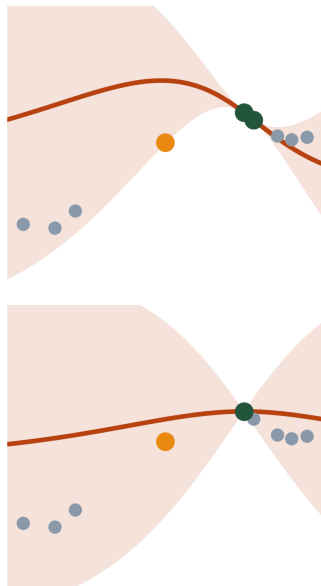Chooses redundant information

Conditional k-th Nearest Neighbors

Naive: select k closest points

Chooses redundant information
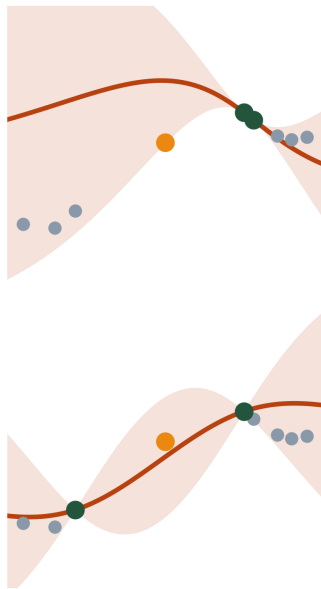
Maximize *mutual information*!

Conditional k-th Nearest Neighbors

Naive: select k closest points

Chooses redundant information
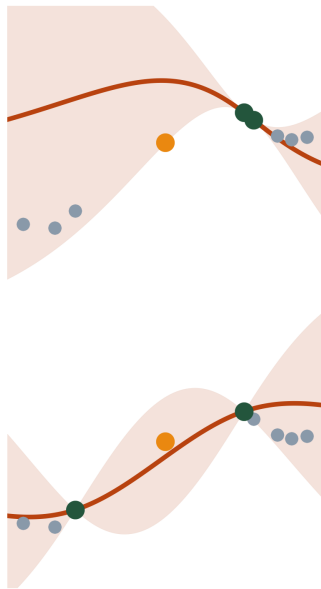
Maximize *mutual information*!

Conditional k-th Nearest Neighbors

Naive: select k closest points

Chooses redundant information

Maximize *mutual information*!

Direct computation: $\mathcal{O}(Nk^4)$

Conditional k-th Nearest Neighbors

Naive: select k closest points

Chooses redundant information

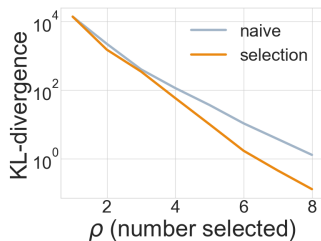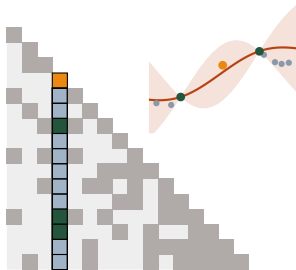Maximize *mutual information*!

Direct computation: $\mathcal{O}(Nk^4)$

Store Cholesky factor $\rightarrow \mathcal{O}(Nk^2)$!

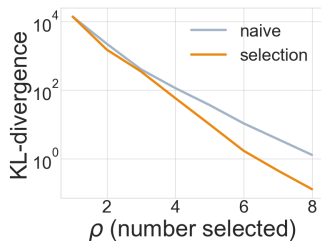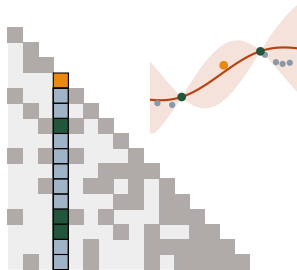# Cholesky Factorization by Selection

Apply column-wise
$\rightarrow$ sparse approx. of GP

Cholesky Factorization by Selection

Apply column-wise
$\rightarrow$ sparse approx. of GP

Maximum mutual information
$\rightarrow$ minimum KL-divergence

Cholesky Factorization by Selection
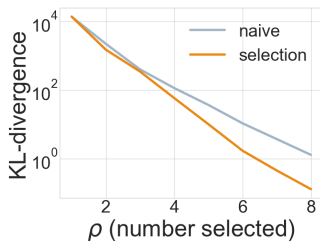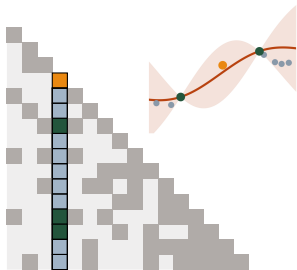

Apply column-wise
$\rightarrow$ sparse approx. of GP

Maximum mutual information
$\rightarrow$ minimum KL-divergence

Improves approx. algorithm of [1]

[1]F. Schäfer, M. Katzfuss, and H. Owhadi, "Sparse Cholesky factorization by Kullback-Leibler minimization," *arXiv preprint arXiv:2004.14455*, 2020