

Moving Least Square Approximation

Hsueh-Ti (Derek) Liu

Introduction

Intuition

- Minimize weighted least square error to get best approximation of data points
- The “weight” is the “similarity” between a points we are considering and other data points
- How to measure similarity is what we need to design

Problem Formulation (Explain in 2D)

Assume data point i : $y_i = z(x_i) + e_i$
data aprox error

Cost function: $\sum_{i=1}^n [y_i - z(x_j)]^2 K(x_i, x_j, y_i, y_j)$

n - number of data points


j - the point we are looking at


K - "Kernel function" which measures points' "similarity"

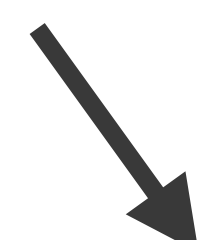
Problem Formulation

$$\min \sum_{i=1}^n [y_i - z(x_j)]^2 K(x_i, x_j, y_i, y_j)$$

$$= [\mathbf{y} - z(x_j)\mathbf{I}_{n \times 1}]^T \mathbf{K} [\mathbf{y} - z(x_j)\mathbf{I}_{n \times 1}]$$


$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$


$$\begin{pmatrix} z_j \\ z_j \\ \vdots \\ z_j \end{pmatrix}$$


$$\begin{pmatrix} K(x_1, x_j, y_1, y_j) & & \\ & K(x_2, x_j, y_2, y_j) & \\ & & \ddots \\ & & & K(x_n, x_j, y_n, y_j) \end{pmatrix}$$

Problem Formulation

Apply solution of weighted least square

$$y = X\beta + \varepsilon$$
$$s = \sum_{i=1}^n w_i (y_i - \hat{y}_i)^2 \quad \rightarrow \quad \hat{\beta} = (X^T W X)^{-1} X^T W y$$

Our solution:

$$\hat{z}(x_j) = (\mathbf{I}^T \mathbf{K} \mathbf{I})^{-1} \mathbf{I}^T \mathbf{K} \mathbf{y}$$
$$= \sum_i \frac{K(x_i, x_j, y_i, y_j)}{\sum_i K(x_i, x_j, y_i, y_j)} y_i$$

Result Intuition

- Approximated data points are just summation of data points with different weights

Kernel Functions

Gaussian Linear function

$$K(x_i, x_j, y_i, y_j) = \exp\left(\frac{-||x_i - x_j||^2}{\sigma^2}\right)$$

- Using geometric distance to penalize outliers
- "sigma" can control how smooth the fitting result is

Other Kernel Functions

- Bilateral Filter
- Non-local Means
- Bilateral Flow Kernel
- LARK Steering Kernel
- ...