GAMES 102 - 作业 1

彭博

November 14, 2020

1. 本次作业使用 4 种不同方法绘制曲线如 Fig.1所示，其中红色曲线为 Lagrange 插值，绿色曲线为径向基函数（RBF）插值，深蓝色曲线为幂函数回归，浅蓝色曲线为岭回归，不同方法的参数可通过上方滚动条进行调节。

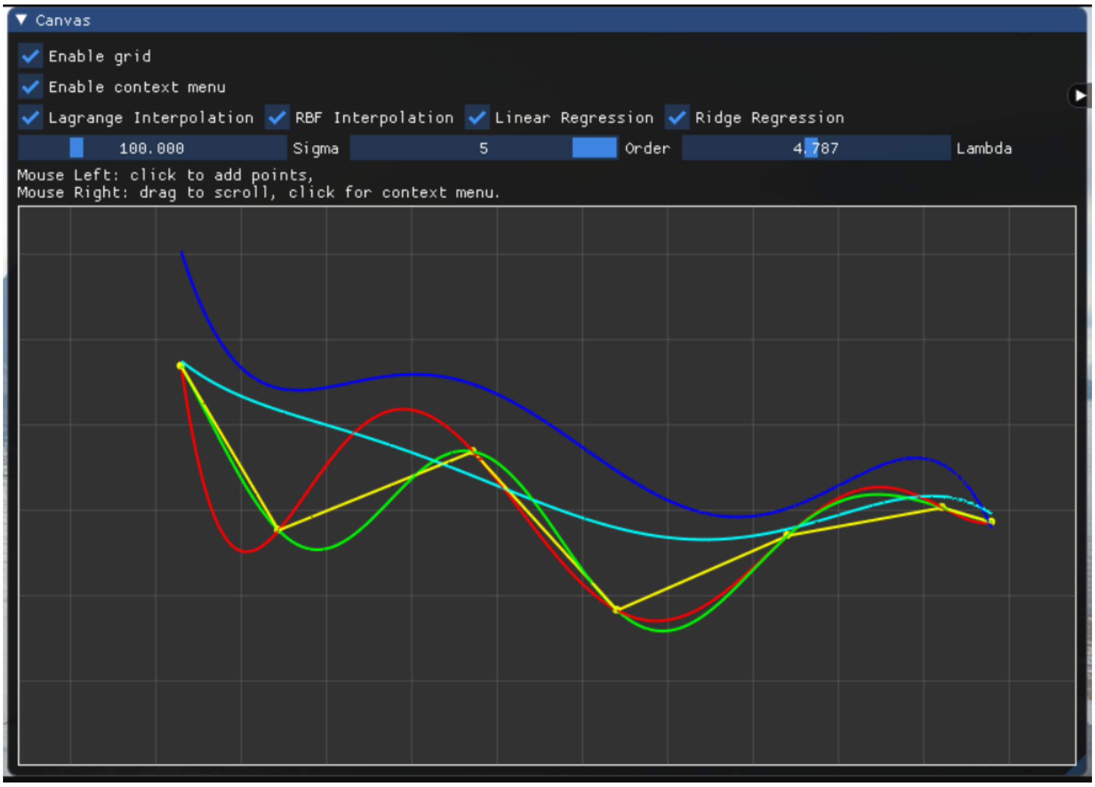

Figure 1: 绘制曲线

不同方法实现原理如下：

(a) Lagrange 插值:
对任意点 (x, y) 其纵坐标可通过 Lagrange 多项式来确定:

$$y = L(x) = \sum_{i=1}^{n} y_i l_i(x) \tag{1}$$

$$l_i(x) = \prod_{j=1, j\neq i}^{n} \frac{x - x_j}{x_i - x_j} \tag{2}$$

其中 (x_i, y_i) 为已知控制点坐标，$l_i(x)$ 为 Lagrange 多项式。

(b) RBF 插值:
对任意点 (x, y) 其纵坐标可通过 RBF 的线性组合来确定:

$$y = \sum_{i=1}^{n} \lambda_i \phi_i(x) \tag{3}$$

$$\phi_i(x) = \exp\{-\frac{(x - x_i)^2}{2\sigma^2}\} \tag{4}$$

其中 (x_i, y_i) 为已知控制点坐标, σ 为参数。将 n 个控制点坐标带入插值公式得到线性方程组:

$$\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \dots & \phi_n(x_n) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{5}$$

求解方程组 $\mathbf{\Phi}\boldsymbol{\lambda} = \mathbf{y}$ 即可获得系数向量 $\boldsymbol{\lambda}$。本次作业中除 RBF 核函数外还增加了常数项 C，使得插值函数变为:

$$y = \sum_{i=1}^{n} \lambda_i \phi_i(x) + C \tag{6}$$

此时待求解变量数 $(n+1)$ 多于方程数 (n)，因此需要引入额外的约束条件:

$$\sum_{i=1}^{n} \lambda_i = 0 \tag{7}$$

重新构造线性方程组得到:

$$\begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots & \phi_n(x_1) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_1(x_n) & \phi_2(x_n) & \dots & \phi_n(x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ C \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \\ 0 \end{bmatrix} \tag{8}$$

重新求解线性方程组即可获得全部系数。
除此之外还可以使用低阶多项式来代替常数项 C，此时插值函数为:

$$y = \sum_{i=1}^{n} \lambda_i \phi_i(x) + \sum_{j=1}^{m} \gamma_j p_j(x) \tag{9}$$

对应的约束条件和线性方程组分别为:

$$\sum_{i=1}^{n} \lambda_i p_j(x_i) = 0 \tag{10}$$

$$\begin{bmatrix} \mathbf{\Phi} & \boldsymbol{P} \\ \boldsymbol{P^T} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ \mathbf{0} \end{bmatrix} \tag{11}$$

重新求解线性方程组即可。

(c) 幂函数回归:
假设曲线可以由一个 k 次多项式来描述:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k \tag{12}$$

对于 n 个控制点可以构造线性方程组 $\boldsymbol{X\beta = y}$:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^k \\ 1 & x_2 & x_2^2 & \dots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^k \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \tag{13}$$

方程数 (n) 多于待求解系数 $(k+1)$，此时采用最小二乘法进行求解，相当于求解正规方程:

$$\boldsymbol{X}^T \boldsymbol{X} \boldsymbol{\beta} = \boldsymbol{X}^T \boldsymbol{y} \tag{14}$$

$\boldsymbol{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$ 即为多项式函数所需系数。

(d) 岭回归:
岭回归与幂函数回归类似，但在构造正规方程时额外添加正则项:

$$(\boldsymbol{X}^T \boldsymbol{X} + \lambda \boldsymbol{I})\boldsymbol{\beta} = \boldsymbol{X}^T \boldsymbol{y} \tag{15}$$

其中，λ 为正则项系数；$\boldsymbol{I}$ 为 $k+1$ 阶单位阵。

对比 4 种不同方法可以发现:

(a) Lagrange 插值和 RBF 插值能够保证曲线通过控制点，而幂函数回归和岭回归在多项式次数较低时不能保证曲线通过控制点；
(b) 控制点较多时使用 Lagrange 插值容易导致曲线发生震荡 (Runge 现象) 如 Fig.2所示；
(c) 使用 RBF 插值时参数 σ 能够控制曲线的光滑程度，σ 过小会导致曲线不够光滑如 Fig.3所示；

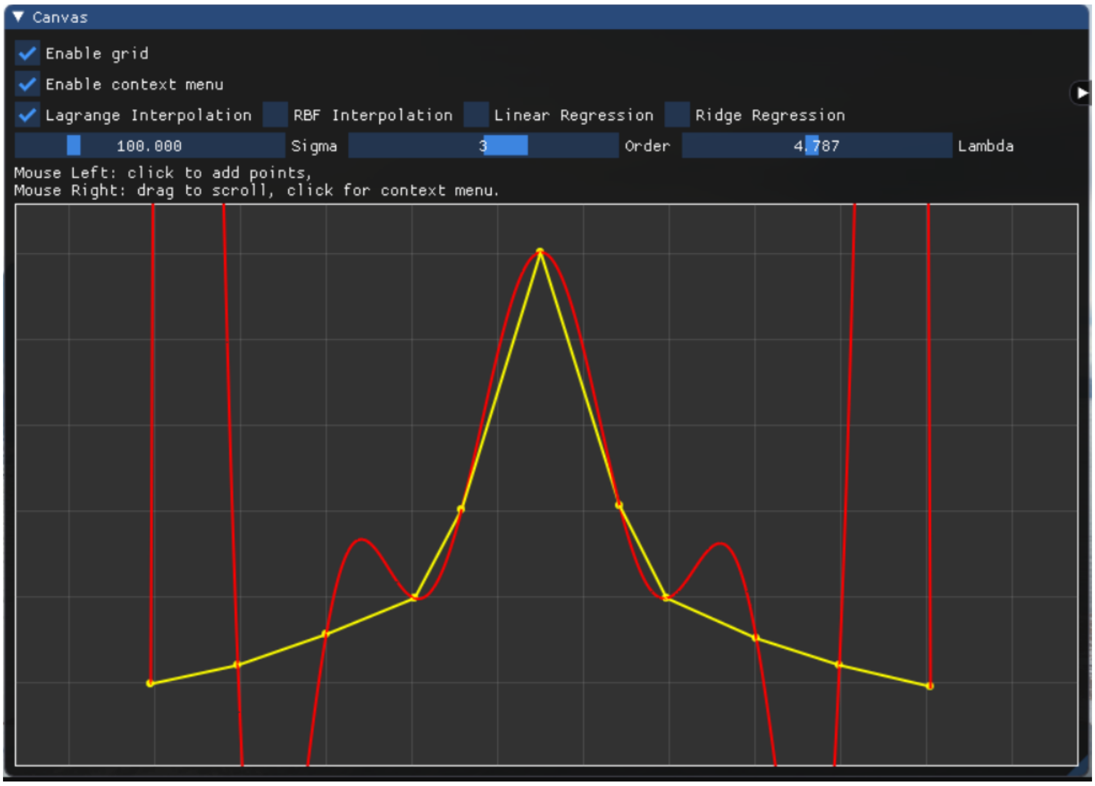

Figure 2: Runge 现象

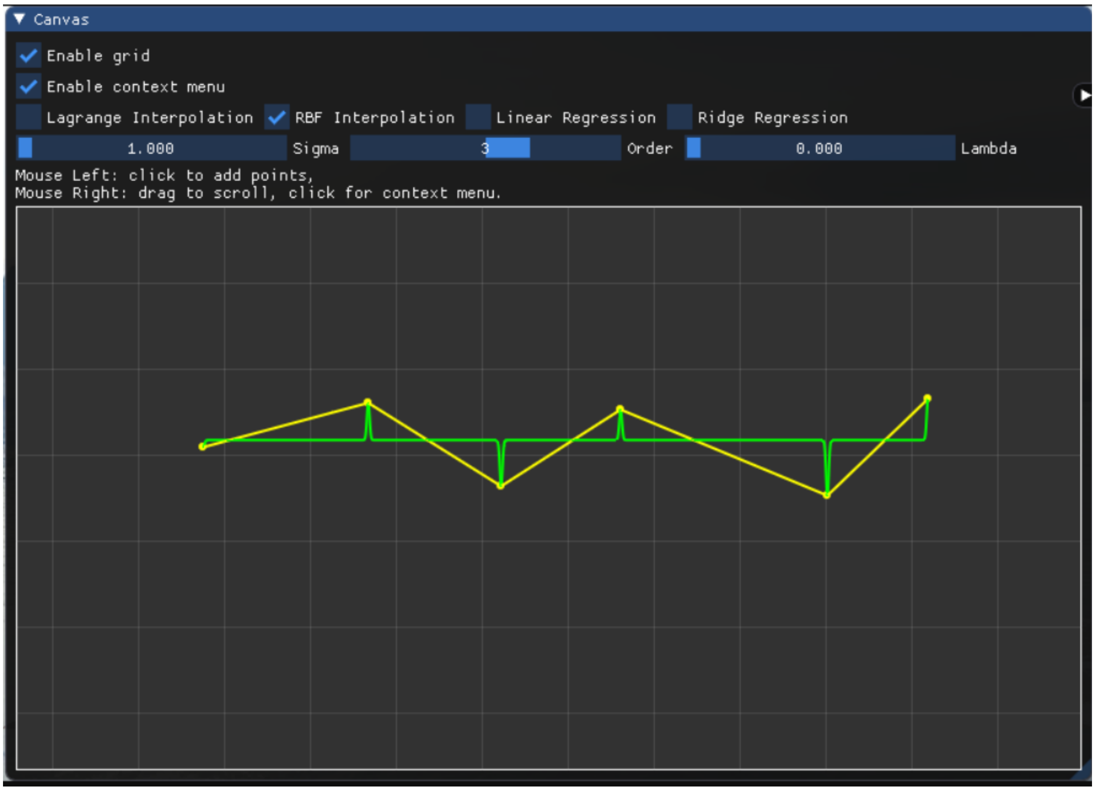

Figure 3: RBF 插值