

地统计系列

地理加权回归模型 -1- 理论推导

UP：小勇啊哈

2023年7月23日

提纲

一、多元线性回归推导

二、为什么会有地理加权回归？

三、地理加权回归推导

多元线性回归推导

一、多元线性回归推导

小勇啊哈 bilibili

多元线性回归 (Multiple Linear Regression)：用于描述一个连续因变量和多个自变量之间的线性依存关系的方法。

数据定义

- 观测（测量）数据集

$$\mathbf{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$$

— n 为样本总数

— \mathbf{x}_i 为第 i 个样本的自变量向量，自变量有 K 个， $\mathbf{x}_i = [1, x_{i1}, x_{i2}, \dots, x_{iK}]^T$ ，其中 1 对应着常数项（非随机部分） $(K+1) \times 1$

— y_i 为第 i 个样本的因变量值

- 待估计数据集

$$\mathbf{D}' = \{\mathbf{x}_0', \mathbf{x}_1', \dots, \mathbf{x}_m'\}$$

问题定义

基于观测数据集，构建线性回归方程 $f(\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i$ ，对带估计数据集 \mathbf{D}' 的 y 进行估计

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_K x_K + \varepsilon_i$$

— $\boldsymbol{\beta}$ 为待估计参数，

$$\boldsymbol{\beta} = [\beta_0, \beta_1, \beta_2, \dots, \beta_K]^T \quad (K+1) \times 1$$

非随机部分 已知信息 随机部分

— ε_i 为随机误差，服从正态分布

$$\varepsilon_i \sim N(0, \sigma^2), \quad \sigma^2 \text{ 为方差}$$

一、多元线性回归推导

小勇啊哈 bilibili

最小二乘法

最优化问题：

$$\hat{\beta} = \min_{\beta} \sum_{i=1}^n (y_i - x_i^T \beta)^2$$

使拟合出的曲线与观测值的平方差最小
即随机误差为0的情况下，让 $f(x_i)$ 尽可能接近 y_i

$$= \min_{\beta} (Y - X\beta)^T (Y - X\beta)$$

矩阵形式表达

$$= \min_{\beta} (Y^T Y - Y^T X\beta - \beta^T X^T Y + \beta^T X^T X\beta)$$

其中：

$$Y = [y_1, y_2, \dots, y_n]^T \quad n \times 1$$

$$X = [x_1, x_2, \dots, x_n]^T \quad n \times (K+1)$$

参数估计：



求偏导

$$\frac{\partial}{\partial \beta}$$

$$-2X^T Y + 2X^T X\hat{\beta} = 0$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

矩阵求导公式参考：

$$\frac{\partial x^T a}{x} = \frac{\partial a^T x}{x} = a$$

$$\frac{\partial x^T x}{x} = 2x \quad \frac{\partial x^T A x}{x} = Ax + A^T x$$

回归求解：

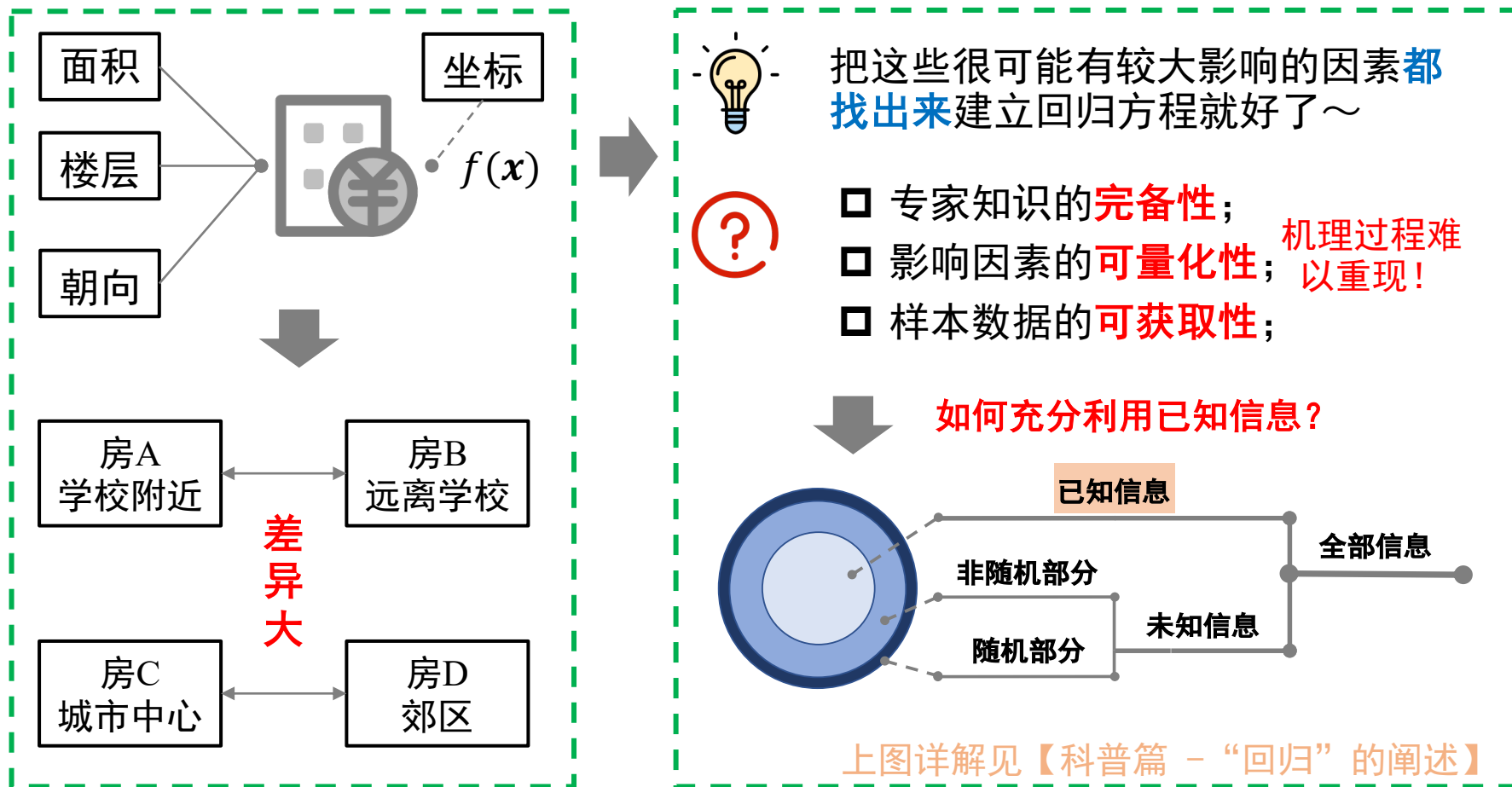
$$y'_0 = x_0'^T \hat{\beta} = x_0'^T (X^T X)^{-1} X^T Y$$

**为什么会有
地理加权回归？**

二、为什么会有地理加权回归？

小勇啊哈 bilibili

以**多元线性回归**得到的全局回归方程，在涉及与**空间位置相关**的回归问题时，常难以表达X与Y之间的关系（即存在较大误差，随机误差 ε 影响过大）。



二、为什么会有地理加权回归？

小勇啊哈 bilibili

较为完备的机理重建往往“**此路不通**”，地理学家通过分析地理现象在空间上的复杂规律，提出**地理学第一定律**和**地理学第二定律**，为回归问题从**空间视角**认知带来了新的机遇，地理加权回归应运而生。



地理学第一定律

任何事物都与其他事物相关，
但是近处的事物比远处的事物更相关。

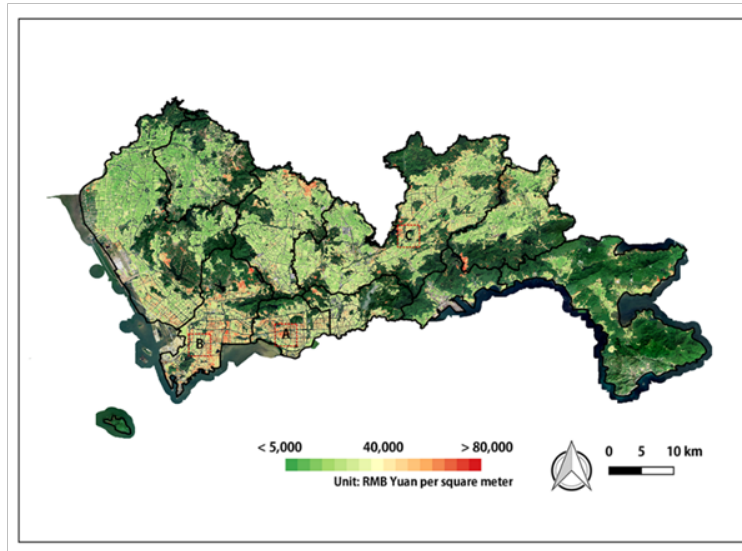
——Waldo R. Tobler



地理学第二定律

地理现象的空间变化以及变化的差异性，即不可控的空间变化规律。

——Michael F. Goodchild



深圳市房价空间分布（姚尧等，2018）

- ✓ 房价分布呈现空间邻近相似；
- ✓ 房价分布难以用全局回归方程概括；

地理加权回归推导

三、地理加权回归推导

小勇啊哈 bilibili

英国Fotheringham教授在1996年便提出**地理加权回归模型**(Geographical Weighted Regression, GWR), GWR旨在区域每一处生成**局部**关系的回归模型。

数据定义

- 观测（测量）数据集

$$\mathbf{D} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$$

样本 (\mathbf{x}_i, y_i) 对应的样本空间位置为 (u_i, v_i)

— n 为样本总数

— \mathbf{x}_i 为第 i 个样本的自变量向量 $(K+1) \times 1$

— y_i 为第 i 个样本的因变量值

- 待估计数据集

$$\mathbf{D}' = \{\mathbf{x}_0', \mathbf{x}_1', \dots, \mathbf{x}_m'\}$$

样本 \mathbf{x}_i' 对应的样本空间位置为 (u_i', v_i')

问题定义

基于观测数据集，构建线性回归方程 $f(\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta}(u_i', v_i') + \varepsilon_i$ ，对带估计数据集 \mathbf{D}' 中 \mathbf{x}_i' 对应的 y_i' 进行估计

为每一个待估计样本计算一组回归参数，即地理学第二定律的体现

— $\boldsymbol{\beta}(u_i', v_i')$ 为样本 \mathbf{x}_i' 的待估计参数，
 $\boldsymbol{\beta}(u_i', v_i') = [\beta_0(u_i', v_i'), \beta_1(u_i', v_i'), \beta_2(u_i', v_i') \dots, \beta_K(u_i', v_i')]^T$
 $(K+1) \times 1$

— ε_i 为随机误差，服从正态分布
 $\varepsilon_i \sim N(0, \sigma^2)$, σ^2 为方差

三、地理加权回归推导

小勇啊哈 bilibili

最小二乘法

最优化问题：【 $\beta(u'_0, v'_0)$ 简化为 β_0 表示】

$$\widehat{\beta}_0 = \min_{\beta} \sum_{i=1}^n w_{i0} [(y_i - \mathbf{x}_i^T \beta_0)]^2$$

$$= \min_{\beta_0} (\mathbf{Y} - \mathbf{X}\beta_0)^T \mathbf{W}_0 (\mathbf{Y} - \mathbf{X}\beta_0)$$

$$= \min_{\beta_0} (\mathbf{Y}^T \mathbf{W}_0 \mathbf{Y} - \mathbf{Y}^T \mathbf{W}_0 \mathbf{X} \beta_0 - \beta_0^T \mathbf{X}^T \mathbf{W}_0 \mathbf{Y} + \beta_0^T \mathbf{X}^T \mathbf{W}_0 \mathbf{X} \beta_0)$$

引入空间权重 w_{i0} ,代表第 i 个观测样本对第0个待估计样本的影响程度,即地理学第一定律体现

矩阵形式表达

参数估计：

求偏导

$$\frac{\partial}{\partial \beta_0}$$

$$-2\mathbf{X}^T \mathbf{W}_0 \mathbf{Y} + 2\mathbf{X}^T \mathbf{W}_0 \mathbf{X} \widehat{\beta}_0 = 0$$

$$\widehat{\beta}_0 = (\mathbf{X}^T \mathbf{W}_0 \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_0 \mathbf{Y}$$

回归求解：

$$y'_0 = \mathbf{x}_0'^T \widehat{\beta}_0 = \mathbf{x}_0'^T (\mathbf{X}^T \mathbf{W}_0 \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}_0 \mathbf{Y}$$

其中：

$$\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$$

$n \times 1$

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^T$$

$n \times (K+1)$

$$\mathbf{W}_0 = \begin{bmatrix} w_{10} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_{n0} \end{bmatrix}$$

$n \times n$

\mathbf{W}_0 为对角矩阵,每个元素的取值范围均为 $[0,1]$

Thanks

空间权重的讲解敬请期待下一个视频.....