

Nonparametric Regression and Polynomial Regression in R

Table of Contents

- [Nonparametric Regression and Polynomial Regression in R](#)
 - **Summary of Key Concepts**
 - **R Code Examples**
 - 1. Loading Data in R
 - 2. Polynomial Regression
 - 3. Nonparametric Regression (Kernel Smoothing)
 - 4. Visualization Fine-tuning
 - **中文知识点总结与R语言实例**
 - 1. 在R中读取数据
 - 2. 多项式回归
 - 3. 非参数回归（核平滑）
 - 4. 图形美化

Summary of Key Concepts

This video discusses the following key concepts in the context of analyzing bone density data:

1. Loading Data in R:

- The video explains how to load bone density data from a specific R package, which is associated with a well-known statistical learning textbook.

2. Polynomial Regression:

- Polynomial regression involves fitting a model where the predictors are raised to different powers (e.g., squared, cubed).
- The video describes two methods for selecting polynomial terms:
 - **Method 1:** Start with a linear model (degree = 1) and add higher-order terms until the new term is no longer statistically significant.
 - **Method 2:** Start with a higher-order polynomial (e.g., degree = 5) and remove terms until only significant ones remain.
- The video concludes that neither method has strong statistical justification and suggests nonparametric regression as an alternative.

3. Nonparametric Regression (Kernel Estimation):

- Nonparametric regression, such as kernel smoothing, allows the shape of the relationship between the predictor and response variables to be learned from the data without assuming a specific form.
- The `ksmooth` function in R is introduced for kernel estimation, where the choice of bandwidth affects the smoothness of the curve.

4. Visualization:

- The video demonstrates how to plot the results of both polynomial regression and kernel smoothing using base R plotting functions.
- It emphasizes the impact of bandwidth on the smoothness of the kernel estimator and discusses how to fine-tune the visual presentation of plots.

R Code Examples

Here are the corresponding R code examples for each concept discussed:

1. Loading Data in R

```
# Load the required package
library(ISLR) # Assuming the bone density data is in the ISLR package

# Load the bone density data
data(BoneDensity)
```

2. Polynomial Regression

Method 1: Adding Terms Until Insignificance

```
# Linear model (degree = 1)
model1 <- lm(BoneDensity ~ Age, data = BoneDensity)
summary(model1)

# Quadratic model (degree = 2)
model2 <- lm(BoneDensity ~ Age + I(Age^2), data = BoneDensity)
summary(model2)
```

Method 2: Removing Terms from Higher-Order Polynomial

```
# Higher-order polynomial (degree = 5)
model5 <- lm(BoneDensity ~ poly(Age, 5), data = BoneDensity)
summary(model5)

# Remove the highest degree term and fit degree 4
model4 <- lm(BoneDensity ~ poly(Age, 4), data = BoneDensity)
summary(model4)
```

3. Nonparametric Regression (Kernel Smoothing)

```
# Kernel smoothing using ksmooth
ksmooth_result <- ksmooth(x = BoneDensity$Age, y = BoneDensity$BoneDensity, kernel = "normal", lwd = 2)

# Plotting the result
plot(BoneDensity$Age, BoneDensity$BoneDensity, pch = 16, col = rgb(0.5, 0.5, 0.5, 0.9))
lines(ksmooth_result$x, ksmooth_result$y, col = "black", lwd = 2)
```

4. Visualization Fine-tuning

```
# Customize plot appearance
plot(BoneDensity$Age, BoneDensity$BoneDensity, pch = 19, col = rgb(0.5, 0.5, 0.5, 0.7), cex = 1.2)
lines(ksmooth_result$x, ksmooth_result$y, col = "black", lwd = 2)
```

中文知识点总结与R语言实例

1. 在R中读取数据

```
# 加载所需的包
library(ISLR) # 假设骨密度数据在ISLR包中

# 读取骨密度数据
data(BoneDensity)
```

2. 多项式回归

方法1: 添加项直到不显著

```
# 线性模型 (degree = 1)
model1 <- lm(BoneDensity ~ Age, data = BoneDensity)
summary(model1)

# 二次模型 (degree = 2)
model2 <- lm(BoneDensity ~ Age + I(Age^2), data = BoneDensity)
summary(model2)
```

方法2: 从高阶多项式中删除项

```
# 高阶多项式 (degree = 5)
model5 <- lm(BoneDensity ~ poly(Age, 5), data = BoneDensity)
summary(model5)

# 移除最高阶项, 拟合四次模型
model4 <- lm(BoneDensity ~ poly(Age, 4), data = BoneDensity)
summary(model4)
```

3. 非参数回归 (核平滑)

```
# 使用ksmooth进行核平滑
ksmooth_result <- ksmooth(x = BoneDensity$Age, y = BoneDensity$BoneDensity, kernel = "normal", lwd = 2)

# 绘制结果
plot(BoneDensity$Age, BoneDensity$BoneDensity, pch = 16, col = rgb(0.5, 0.5, 0.5, 0.9))
lines(ksmooth_result$x, ksmooth_result$y, col = "black", lwd = 2)
```

4. 图形美化

```
# 自定义图形外观
plot(BoneDensity$Age, BoneDensity$BoneDensity, pch = 19, col = rgb(0.5, 0.5, 0.5, 0.7), cex = 1.2)
lines(ksmooth_result$x, ksmooth_result$y, col = "black", lwd = 2)
```

These summaries and examples should help you better understand the concepts discussed in the video and how to implement them in R.