STAT40810 — Stochastic Models

Brendan Murphy

Week 4

Spline Smoothing

Spline Smoothing: Motivation

- We want to estimate the function $f(x)$ where $y_i = f(x_i) + \epsilon_i$.
- In principle, we could choose knot locations and a spline order (typically cubic) and find the least squares fit.

Spline Smoothing: Motivation

- We want to estimate the function $f(x)$ where $y_i = f(x_i) + \epsilon_i$.
- In principle, we could choose knot locations and a spline order (typically cubic) and find the least squares fit.
- Alternatively, we could try to estimate $f(x)$ without putting many restrictions on its functional form.
- Minimizing $\sum_{i=1}^{n}[y_i - f(x_i)]^2$ would give a well fitting model. However, without restrictions, the $f(x)$ that minimizes this criterion interpolates the data (unless the x_i values are non-unique).

Spline Smoothing: Motivation

- We want to estimate the function $f(x)$ where $y_i = f(x_i) + \epsilon_i$.
- In principle, we could choose knot locations and a spline order (typically cubic) and find the least squares fit.
- Alternatively, we could try to estimate $f(x)$ without putting many restrictions on its functional form.
- Minimizing $\sum_{i=1}^{n}[y_i - f(x_i)]^2$ would give a well fitting model. However, without restrictions, the $f(x)$ that minimizes this criterion interpolates the data (unless the x_i values are non-unique).
- If we put a penalty for the "roughness of our estimate", then we may get a more sensible result.

Roughness Penalty

- One possibility is to penalize functions that are two "rough" or "wiggly".
- One criterion for measuring roughness is,

$$R(f^{(j)}) = \int [f^{(j)}(x)]^2 dx. \text{ } (f^{(j)} \text{ is the } j\text{th derivative}).$$

Roughness Penalty

- One possibility is to penalize functions that are two "rough" or "wiggly".
- One criterion for measuring roughness is,

$$R(f^{(j)}) = \int [f^{(j)}(x)]^2 dx. \ (f^{(j)} \text{ is the } j\text{th derivative}).$$

- It is worth noting that if

$$f(x) = \beta_0 + \sum_{p=1}^{P} \beta_p x^p,$$

that is that $f(x)$ is a polynomial of order P.
Then,

$$R(f^{(Q)}) = 0 \text{ for } Q > P$$

Spline Smoothing

- Suppose we want to find the $f(x)$ that minimizes

$$\sum_{i=1}^{n}[y_i - f(x_i)]^2 + \lambda \int [f^{(j)}(x)]^2 dx.$$

- This criterion balances the goodness of fit of f with the roughness of f.

[1] To be exact it's something called a natural spline.

Spline Smoothing

- Suppose we want to find the $f(x)$ that minimizes

$$\sum_{i=1}^{n}[y_i - f(x_i)]^2 + \lambda \int [f^{(j)}(x)]^2 dx.$$

- This criterion balances the goodness of fit of f with the roughness of f.
- The function that minimizes this criterion is a function that is polynomial of order $(j+1)$ between distinct x values and has j continuous derivatives.
- In other words... a spline.[1]

[1] To be exact it's something called a natural spline.

Spline Smoothing

- Suppose we want to find the $f(x)$ that minimizes

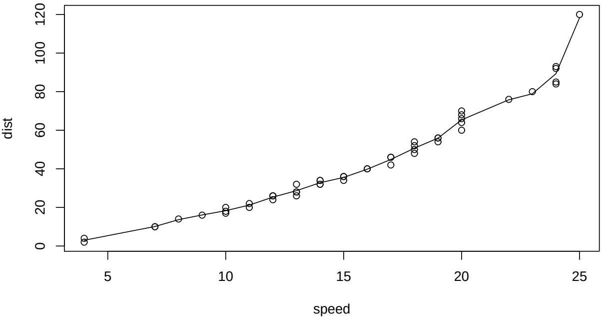
$$\sum_{i=1}^{n}[y_i - f(x_i)]^2 + \lambda \int [f^{(j)}(x)]^2 dx.$$

- This criterion balances the goodness of fit of f with the roughness of f.
- The function that minimizes this criterion is a function that is polynomial of order $(j+1)$ between distinct x values and has j continuous derivatives.
- In other words... a spline.[1]
- The case of most interest is when $j = 2$ and this gives cubic splines.

[1] To be exact it's something called a natural spline.

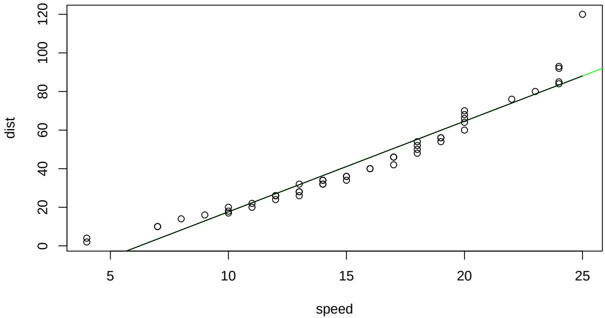
Example: Breaking Distance

- Suppose that we minimize the criterion with $\lambda = 0.2$, then we get a flexible fit.



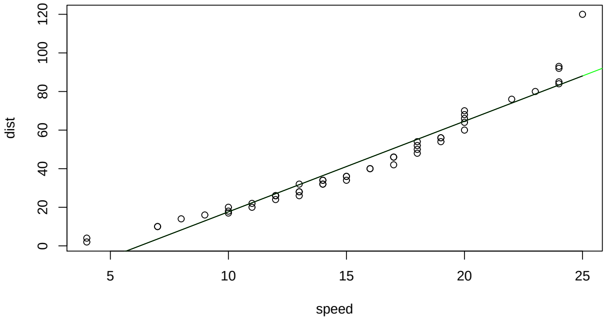
Example: Breaking Distance

- Suppose that we minimize the criterion with $\lambda = 1.5$, then we get a curve that's almost a linear fit.



Example: Breaking Distance

- Suppose that we minimize the criterion with $\lambda = 1.5$, then we get a curve that's almost a linear fit.



- The choice of λ drives the balance between fit and smoothness.