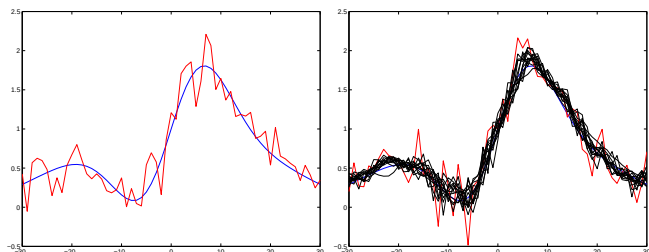


Stat 471: Lecture 29

Monte-Carlo smoothing.

Moo K. Chung
mchung@stat.wisc.edu

November 17, 2003



This is one example of assigning probabilities to each y_i . We estimate g as

$$\hat{g}(p_0) = \mathbb{E}_w Y_0 = \sum_{i=0}^m w_i y_i. \quad (2)$$

Note that when $p_1 = p_2 = \dots = p_m$, (1) becomes a special case of (2). The estimation (2) is the *multivariate kernel smoothing* we studied in lecture 28.

1. In the previous lecture, we introduced the *deterministic* bivariate smoothing. There is a Monte-Carlo version of bivariate smoothing. Suppose we have a following model

$$Y = g(p) + \epsilon$$

where $p_i \in \mathbb{R}^k$. For example, we might have the following simple linear model

$$\text{housevalue} = \beta_0 + \beta_1 \text{longitude} + \beta_2 \text{latitude} + \text{error}.$$

Assuming $\mathbb{E}\epsilon = 0$, we have $g(p) = \mathbb{E}Y$. If we have m observations at position p_0 , we might have m different observations $y_0^1, \dots, y_0^m$. So we can estimate regression function g at position p_0 as

$$\hat{g}(p_0) = \frac{1}{m} \sum_{j=1}^m y_0^j. \quad (1)$$

2. However, we usually do not have repeated measurements at position p_0 . Rather we will have observations at $p_1, \dots, p_m$ around p_0 . Let $y_1, \dots, y_m$ be the observations taken at $p_1, \dots, p_m$ respectively. If $\|p_i - p_0\|$ is small, we can take y_i to be another measurement at p_0 . On the other hand if $\|p_i - p_0\|$ is large, we may assume that observation y_i is highly unlikely to be another measurement at p_0 . Then we can assign probability to each observation y_i at p_i .

$$w_i = P(y_i \text{ will occur at } p_0) \propto \exp\left(-\frac{\|p_i - p_0\|^2}{2\sigma^2}\right).$$

We normalize it such that

$$w_i = \frac{\exp\left(-\frac{\|p_i - p_0\|^2}{2\sigma^2}\right)}{\sum_{i=0}^m \exp\left(-\frac{\|p_i - p_0\|^2}{2\sigma^2}\right)}.$$

3. Instead of computing weights directly, there is an alternate Monte-Carlo method. Suppose $X \sim f$ be the density of X . Then for function $Y(p) = g(p) + \epsilon(p)$,

$$\begin{aligned} \mathbb{E} Y(X) &= \frac{\int y(x) f(x) dx}{\int f(x) dx} \\ &\approx \sum_{i=0}^m w_i y(p_i) = \sum_{i=0}^m w_i y_i. \end{aligned}$$

We will choose f such that it will give weights w_i similar to the ones described in the previous section. Then we estimate

$$\hat{g}(p_0) = \frac{1}{n} \sum_{i=1}^n y(X_i),$$

where $p(X_i = p_i) = w_i$.

```
x=-30:30;
g=0.2*x.*exp(-x.^2/100) + cos(x/30).^2;
y=g+0.2*randn(1,61);
```

```
m=10; sigma=2;
temp_y=y;
mc_y=zeros(1,61); mc_run=50;
for i=1:mc_run
    for i=1:61
        X=fix(normrnd(i,sigma,100,1));
        p=X(find((X>=1) & (X <=61)));
        p=p(1:m);
        smooth_y(i) = mean(temp_y(p));
    end;
    %hold on;plot(x,smooth_y,'k:')
    mc_y=mc_y+smooth_y;
end;
mc_y=mc_y/mc_run;
```