Parameter Estimation

Raghu Pasupathy (Virginia Tech) and Shane G. Henderson (Cornell)

May 29, 2007

(Continuous variables, unconstrained, unknown if it is smooth or not.)

Problem Statement: Say a simulation generates output data $\{Y_j\}, Y_j \in [0, \infty] \times [0, \infty]$, that are i.i.d and known to come from a distribution with the two-dimensional density function

$$f(y_1, y_2; x^*) = \frac{e^{-y_1} y_1^{x_1^* y_2 - 1}}{\Gamma(x_1^* y_2)} \frac{e^{-y_2} y_2^{x_2^* - 1}}{\Gamma(x_2^*)}, \quad y_1, y_2 > 0,$$

where $x^* \equiv (x_1^*, x_2^*)$ is the unknown vector of parameters.

Noting that x^* maximizes the function

$$\begin{aligned} g(x) &= \mathrm{E}\left[\log\left(f(Y; x)\right)\right] \\ &= \int_0^\infty \log\left(f(y; x)\right) f(y; x^*) dy, \end{aligned}$$

and that

$$G_m(x) = \frac{\sum_{j=1}^m \log(f(Y_j; x))}{m}$$

is a consistent estimator of $g(x)$, find x^*.

Recommended Parameter Settings: Use $x^* = (2, 5)$ to generate the data. A poor objective function value is the true objective function value associated with taking $x = (1, 1)$.

Starting Solution(s): Take $x = (1, 1)$. If multiple solutions are required then select them as i.i.d. uniform in the open square $(0, 10) \times (0, 10)$.

Measurement of time: Number of output data points generated. Take as budget $1000, 10000$.

Optimal Solution(s): Global maximum at $x^* = (2, 5)$.

Known Structure: None.