M3/M4S3 STATISTICAL THEORY II

TWO USEFUL RESULTS

Notation: First Derivatives

Suppose $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ is a real function, and denote the $(1 \times k)$ vector of *first* partial derivatives by $\dot{\boldsymbol{f}}$, that is, for $\boldsymbol{x} \in \mathbb{R}^k$,

$$\dot{\boldsymbol{f}}(\boldsymbol{x}) = \left[\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \ldots, \frac{\partial f(\boldsymbol{x})}{\partial x_k}\right]$$

By extension, if $\boldsymbol{f} : \mathbb{R}^k \longrightarrow \mathbb{R}^d$ is a real function, then we regard it as a $d \times 1$ vector

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_d(\boldsymbol{x}) \end{bmatrix}$$

and define the $(d \times k)$ matrix of first partial derivatives with $(j, l)^{\text{th}}$ element

$$\frac{\partial f_j(\boldsymbol{x})}{\partial x_l},$$

that is,

$$\dot{\boldsymbol{f}}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \dfrac{\partial f_1(\boldsymbol{x})}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_d(\boldsymbol{x})}{\partial x_1} & \cdots & \dfrac{\partial f_d(\boldsymbol{x})}{\partial x_k} \end{bmatrix}$$

Second Derivatives

Suppose $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ is a real function, we define the $(k \times k)$ matrix of *second* partial derivatives, $\ddot{f}(\boldsymbol{x})$, with $(j, l)^{\text{th}}$ element

$$\frac{\partial^2}{\partial x_j \partial x_l} f(\boldsymbol{x}),$$

that is,

$$\ddot{f}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_k} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_k \partial x_1} & \cdots & \dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_k^2} \end{bmatrix}$$

The Mean-Value Theorem
Suppose that $\boldsymbol{f} : \mathbb{R}^k \longrightarrow \mathbb{R}^d$ is a real function, and that $\dot{\boldsymbol{f}}(\boldsymbol{x})$ is continuous in the ball of radius $r > 0$ centered at $\boldsymbol{x_0}$, that is, in

$$\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x_0}\| < r\}.$$

Then for $\|\boldsymbol{t}\| < r$, $\boldsymbol{t} \in \mathbb{R}^k$,

$$\boldsymbol{f}(\boldsymbol{x_0} + \boldsymbol{t}) = \boldsymbol{f}(\boldsymbol{x_0}) + \left\{\int_0^1 \dot{\boldsymbol{f}}(\boldsymbol{x_0} + u\boldsymbol{t})\, du\right\} \boldsymbol{t}$$

Proof. Let $\boldsymbol{h}(u) = \boldsymbol{f}(\boldsymbol{x_0} + u\boldsymbol{t})$, so that, by the chain rule, $\dot{\boldsymbol{h}}(u) = \dot{\boldsymbol{f}}(\boldsymbol{x_0} + u\boldsymbol{t})\boldsymbol{t}$. Then

$$\left\{\int_0^1 \dot{\boldsymbol{f}}(\boldsymbol{x_0} + u\boldsymbol{t})\, du\right\} \boldsymbol{t} = \int_0^1 \dot{\boldsymbol{h}}(u)\, du = \boldsymbol{h}(1) - \boldsymbol{h}(0) = \boldsymbol{f}(\boldsymbol{x_0} + \boldsymbol{t}) - \boldsymbol{f}(\boldsymbol{x_0}).$$

and the result follows.

Taylor's Theorem
Suppose that $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ is a real function, and that $\ddot{f}(\boldsymbol{x})$ is continuous in the ball of radius $r > 0$ centered at $\boldsymbol{x_0}$, that is, in

$$\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x_0}\| < r\}.$$

Then for $\|\boldsymbol{t}\| < r$, $\boldsymbol{t} \in \mathbb{R}^k$,

$$f(\boldsymbol{x_0} + \boldsymbol{t}) = f(\boldsymbol{x_0}) + \dot{f}(\boldsymbol{x_0})\boldsymbol{t} + \boldsymbol{t}^{\mathsf{T}} \left\{\int_0^1 \int_0^1 v\ddot{f}(\boldsymbol{x_0} + uv\boldsymbol{t})\, du dv\right\} \boldsymbol{t}$$

Note that in these two results,

$$\int_0^1 \dot{\boldsymbol{f}}(\boldsymbol{x_0} + u\boldsymbol{t})\, du$$

is a $d \times k$ matrix, and

$$\int_0^1 \int_0^1 v\ddot{f}(\boldsymbol{x_0} + uv\boldsymbol{t})\, du dv$$

is a $k \times k$ matrix.