YOUNG'S THEOREM

A very important and useful result in the calculus of functions of several variables is the following.

> ***Theorem:*** *Let* $f:\mathbb{R}^n \to \mathbb{R}$ *be twice continuously differentiable on its domain of definition,* $\mathbb{X} \subset \mathbb{R}$. *Then on the interior of its domain, the n×n matrix of second-order partial derivatives is symmetric,*
>
> $$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \ \forall \ i, j = 1, \ldots, n.$$

Proof: By holding the x_k's fixed for all $k \neq i, j$ with i and j arbitrary, we can reduce the problem to the simple case of a function of two variables. Therefore, let $y = x_i$, $z = x_j$, and $g(y,z) \equiv f(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_{j-1},z,x_{j+1},\ldots,x_n)$. Note that continuity of the first- and second-order partial derivatives of *f* is equivalent to continuity of the first- and second-order partial derivatives of *g*. Now consider the quantity

(1) $$h(\Delta y,\Delta z) \equiv \left[g(y+\Delta y, z+\Delta z) - g(y, z+\Delta z)\right] - \left[g(y+\Delta y, z) - g(y,z)\right].$$

There are three aspects of the function $h(\Delta y,\Delta z)$ that are important. First, we can interchange the middle two terms of $h(\Delta y,\Delta z)$ without affecting its value to get

(1') $$h(\Delta y,\Delta z) \equiv \left[g(y+\Delta y, z+\Delta z) - g(y+\Delta y, z)\right] - \left[g(y, z+\Delta z) - g(y,z)\right].$$

Second, referring to the right-hand-side of (1), it is clear that for any Δz, $h(0,\Delta z) \equiv 0$. Similarly, referring to the right-hand-side of (1'), it is also clear that for any Δy, $h(\Delta y, 0) \equiv 0$. This means that the graph of $h(\Delta y,\Delta z)$ as a function of Δy (for fixed Δz)

begins at the origin, has base equal to *Δy* and height equal to $h(\Delta y, \Delta z)$, while the graph of $h(\Delta y, \Delta z)$ as a function of *Δz* (for fixed *Δy*) begins at the origin, has base equal to *Δz* and height equal to $h(\Delta y, \Delta z)$. Third, by the composite function theorem, the partial derivatives of $h(\Delta y, \Delta z)$ with respect to *Δy* and *Δz* are

(2) $$\frac{\partial h(\Delta y, \Delta z)}{\partial \Delta y} = \left[\frac{\partial g(y+\Delta y, z+\Delta z)}{\partial y} - \frac{\partial g(y+\Delta y, z)}{\partial y} \right],$$

(2') $$\frac{\partial h(\Delta y, \Delta z)}{\partial \Delta z} = \left[\frac{\partial g(y+\Delta y, z+\Delta z)}{\partial z} - \frac{\partial g(y, z+\Delta z)}{\partial z} \right].$$

Finally, from (2) and (2'), the second-order cross-partial derivatives of $h(\Delta y, \Delta z)$ are

(3) $$\frac{\partial^2 h(\Delta y, \Delta z)}{\partial \Delta y \partial \Delta z} = \frac{\partial^2 g(y+\Delta y, z+\Delta z)}{\partial y \partial z},$$

(3') $$\frac{\partial^2 h(\Delta y, \Delta z)}{\partial \Delta z \partial \Delta y} = \frac{\partial^2 g(y+\Delta y, z+\Delta z)}{\partial z \partial y}.$$

Therefore, by the mean value theorem, for any given *Δz* there is a $c \in [0, \Delta y]$ such that

(4) $$h(\Delta y, \Delta z) = \frac{\partial h(c, \Delta z)}{\partial \Delta y} \cdot \Delta y = \left[\frac{\partial g(y+c, z+\Delta z)}{\partial y} - \frac{\partial g(y+c, z)}{\partial y} \right] \cdot \Delta y.$$

Note that by the hypotheses of the theorem, the terms in square brackets on the right-hand-side of (4) are continuously differentiable and therefore satisfy the conditions required for a second application of the mean value theorem. Therefore, there is a $d \in [0, \Delta z]$ such that

(5) $$h(\Delta y, \Delta z) = \frac{\partial^2 g(y+c, z+d)}{\partial y \partial z} \cdot \Delta y \cdot \Delta z.$$

Next, note that the definition of $h(\Delta y, \Delta z)$ in (1) is "symmetrical" in *y*, *z* and in *Δy*, *Δz*, a fact that lead us to the equivalent expression for $h(\Delta y, \Delta z)$ given in (1'). Therefore, using exactly the same arguments as above, but in the reverse sequence for *Δy* and *Δz* and applied to (1') rather than (1), we obtain

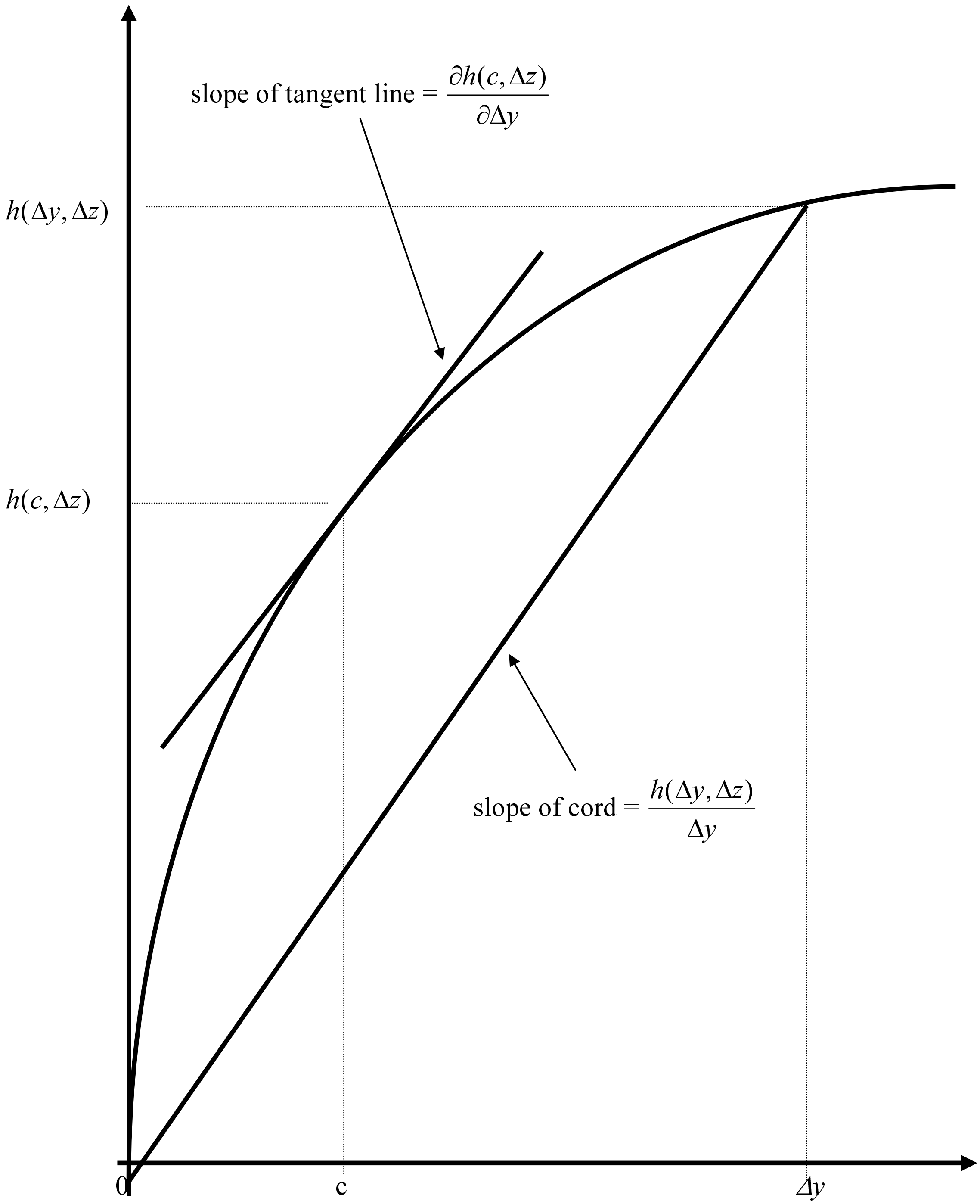
(6) $$h(\Delta y, \Delta z) = \frac{\partial^2 g(y+\tilde{c}, z+\tilde{d})}{\partial z \partial y} \cdot \Delta y \cdot \Delta z$$

for some $\tilde{c} \in [0, \Delta y]$ and some $\tilde{d} \in [0, \Delta z]$. Equating the two expressions (5) and (6) for $h(\Delta y, \Delta z)$ and canceling *Δy* and *Δz* implies

(7) $$\frac{\partial^2 g(y+c, z+d)}{\partial y \partial z} = \frac{\partial^2 g(y+\tilde{c}, z+\tilde{d})}{\partial z \partial y}.$$

As $\Delta y \to 0$ we have $c \to 0$ (since $0 \le |c| \le |\Delta y|$) so that $y + c \to y \equiv x_i$, while as $\Delta z \to 0$ we have $d \to 0$ (since $0 \le |d| \le |\Delta z|$) and $z + d \to z \equiv x_j$. Similarly, as $\Delta y \to 0$ we have $\tilde{c} \to 0$ and $y + \tilde{c} \to y \equiv x_i$, while as $\Delta z \to 0$ we have $\tilde{d} \to 0$ and $z + \tilde{d} \to z \equiv x_j$. Continuity of both sides of (7) then implies that

(8) $$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 g(y,z)}{\partial y \partial z} = \frac{\partial^2 g(y,z)}{\partial z \partial y} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$ **Q.E.D.**



slope of tangent line = $\frac{\partial h(c,\Delta z)}{\partial \Delta y}$
$h(\Delta y, \Delta z)$
$h(c, \Delta z)$
slope of cord = $\frac{h(\Delta y,\Delta z)}{\Delta y}$
0
c
Δy