21 MAY 2013

LINEAR ALG & MULTIVARIABLE CALC

§15

Definition 1. For a function $f\colon \mathcal{D}^n \to \mathbf{R}^m$, the *matrix of partial derivatives of f at the point a* is the matrix $\mathsf{D} f(a)$ whose (i,j) entry is $\frac{\partial f_i}{\partial x_j}(a)$:

$$\mathsf{D} f(a) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_i}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix}$$

Each row corresponds to a component and each column corresponds to a coordinate variable. To check your indexing is correct, make sure that your matrix $\mathsf{D} f(a)$ has the same dimensions as the matrix for a linear transformation $\mathbf{R}^n \to \mathbf{R}^m$. For example, we need the expression $\mathsf{D} f(a)(x-a)$ to make sense. Moreover, we want the resulting vector to be in the codomain $\mathbf{R}^m$.

15.1 Chain Rule

Proposition 1. The matrix of partial derivatives of a composition

$$\begin{gathered} \mathcal{D}^n \xrightarrow{f} (\mathbf{R}^m \supseteq \mathcal{E}^m) \xrightarrow{g} \mathbf{R}^p \\ a \xmapsto{f} f(a) \xmapsto{g} g \circ f(a) \end{gathered}$$

is the composition

$$\begin{gathered} \mathbf{R}^n \xrightarrow{\mathsf{D} f} \mathbf{R}^m \xrightarrow{\mathsf{D} g} \mathbf{R}^p \\ a \xmapsto{f} f(a) \xmapsto{g} g \circ f(a) \end{gathered}$$

which means:

$$\mathsf{D}(g \circ f)(a) = \mathsf{D} g(f(a))\, \mathsf{D} f(a)$$

Example 1 (Licata 8.7). For

$$\begin{aligned} f(x,y,z) &= (\sin x \cos y + e^z, xy \ln(xyz) + xyz^2) \\ g(r,s) &= (1/s, 1/r, s^2) \end{aligned}$$

compute $D(f \circ g)$, the matrix of partial derivatives of the composition $f \circ g$, in two ways. First, write out the composition $f \circ g$ and compute $D(f \circ g)$ directly. Then compute $D(f \circ g)$ using the Chain Rule.

Solution. The composition $f \circ g$ is given by:

$$\begin{aligned} f \circ g(r,s) &= f(1/s, 1/r, s^2) \\ &= \left(\sin \tfrac{1}{s} \cos \tfrac{1}{r} + e^{s^2}, \tfrac{1}{s}\tfrac{1}{r} \ln(\tfrac{1}{s}\tfrac{1}{r}s^2) + \tfrac{1}{s}\tfrac{1}{r}(s^2)^2\right) \\ &= \left(\sin \tfrac{1}{s} \cos \tfrac{1}{r} + e^{s^2}, \tfrac{1}{sr} \ln(\tfrac{s}{r}) + \tfrac{s^3}{r}\right) \end{aligned}$$

Therefore the matrix of partial derivatives of the composition is:

$$D(f \circ g) = \begin{bmatrix} (\sin \frac{1}{s})(-\sin \frac{1}{r})(-\frac{1}{r^2}) & (\cos \frac{1}{s})(-\frac{1}{s^2}) \cos \frac{1}{r} + e^{s^2}(2s) \\ \frac{-1}{sr^2} \ln(\frac{s}{r}) + \frac{1}{sr}(-\frac{1}{r}) - \frac{s^3}{r^2} & \frac{-1}{s^2 r} \ln(\frac{s}{r}) + \frac{1}{sr}(\frac{1}{s}) + \frac{3s^2}{r} \end{bmatrix}$$

Alternatively, we compute the matrix of differentials of each of f and g:

$$D f(x,y,z) = \begin{bmatrix} \cos x \cos y & \sin x(-\sin y) & e^z \\ y \ln(xyz) + \frac{xy}{x} + yz^2 & x \ln(xyz) + \frac{xy}{y} + xz^2 & \frac{xy}{z} + xy(2z) \end{bmatrix}$$

$$D g(r,s) = \begin{bmatrix} 0 & -\frac{1}{s^2} \\ -\frac{1}{r^2} & 0 \\ 0 & 2s \end{bmatrix}$$

and then use the chain rule to write

$$\begin{aligned} D(f \circ g)(r,s) &= D f(g(r,s))\, D g(r,s) \\ &= D f(1/s, 1/r, s^2)\, D g(r,s) \end{aligned}$$

so:

$$D(f \circ g)(r,s) = \begin{bmatrix} \cos \frac{1}{r} \cos \frac{1}{s} & \sin \frac{1}{r}(-\sin \frac{1}{s})) & e^{s^2} \\ \frac{1}{r} \ln(\frac{s}{r}) + \frac{1}{r} + \frac{s^4}{r} & \frac{1}{s} \ln(\frac{s}{r}) + \frac{1}{s} + s^3 & \frac{1}{rs^3} + \frac{2s}{r} \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{s^2} \\ -\frac{1}{r^2} & 0 \\ 0 & 2s \end{bmatrix} \quad \blacksquare$$

Example 2 (Licata 8.13). Assume $\boldsymbol{h}: \mathbf{R}^2 \to \mathbf{R}^3$ satisfies $\boldsymbol{h}(1,2) = (4,5,6)$ and:

$$D \boldsymbol{h}(1,2) = \begin{bmatrix} 0 & 4 \\ 1 & 2 \\ -3 & 6 \end{bmatrix}$$

Define $g: \mathbf{R}^3 \to \mathbf{R}^4$ by $g(r,s,t) = (r, s^2, t^3, -t^2 + 4rs)$. Compute $D(g \circ \boldsymbol{h})(1,2)$.

Solution. By the chain rule:

$$\mathrm{D}(\boldsymbol{g} \circ \boldsymbol{h})(1,2) = \mathrm{D}\,\boldsymbol{g}(\boldsymbol{h}(1,2))\,\mathrm{D}\,\boldsymbol{h}(1,2)$$

Since

$$\mathrm{D}\,\boldsymbol{g}(r,s,t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2s & 0 \\ 0 & 0 & 3t^2 \\ 4s & 4r & -2t \end{bmatrix}$$

and in particular

$$\mathrm{D}\,\boldsymbol{g}(\boldsymbol{h}(1,2)) = \mathrm{D}\,\boldsymbol{g}(4,5,6) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 108 \\ 20 & 16 & -12 \end{bmatrix}$$

so

$$\mathrm{D}(\boldsymbol{g} \circ \boldsymbol{h})(1,2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 108 \\ 20 & 16 & -12 \end{bmatrix} \begin{bmatrix} 0 & 4 \\ 1 & 2 \\ -3 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ 10 & 20 \\ -324 & 648 \\ 52 & 40 \end{bmatrix} \qquad \blacksquare$$

15.2 Linearization/Taylor's Theorem

For $\boldsymbol{x}$ near $\boldsymbol{a}$:

$$f(\boldsymbol{x}) \approx \underbrace{f(\boldsymbol{a}) + \left[\mathrm{D}\, f(\boldsymbol{a})\right](\boldsymbol{x} - \boldsymbol{a})}_{\text{linearization}}$$

$$f(\boldsymbol{x}) \approx \underbrace{f(\boldsymbol{a}) + \left[\mathrm{D}\, f(\boldsymbol{a})\right](\boldsymbol{x} - \boldsymbol{a}) + \tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^{\mathrm{T}}\left[\mathrm{H}\, f(\boldsymbol{a})\right](\boldsymbol{x} - \boldsymbol{a})}_{\text{second degree Taylor polynomial}}$$

Example 3 (Licata 11.7). Compute the linearization of the function

$$q(w,x,y,z) = xy + zw^2 - 3yzw$$

at the point $\boldsymbol{a} = (2, 1, -1, 0)$. ✤

Solution. The matrix of partial derivatives of q is

$$\mathrm{D}\,q(w,x,y,z) = \begin{bmatrix} z(2w) - 3yz & y & x - 3zw & w^2 - 3yw \end{bmatrix}$$

and specializing to $(w,x,y,z) = (2, 1, -1, 0)$ yields:

$$\mathrm{D}\,q(2,1,-1,0) = \begin{bmatrix} 0 & -1 & 1 & 10 \end{bmatrix}$$

Also $q(2,1,-1,0) = -1$. Therefore the linearization of q at $(2,1,-1,0)$ is:

$$L(w,x,y,z) = -1 + \begin{bmatrix} 0 & -1 & 1 & 10 \end{bmatrix} \begin{bmatrix} w-2 \\ x-1 \\ y-(-1) \\ z-0 \end{bmatrix}$$

■

Example 4 (Licata 11.11). Compute the second order Taylor approximation of the function $b(w,z) = w^{3/2} + z^{5/2}$ at $\boldsymbol{a} = (1,4)$. Approximate $b(1.02, 3.96)$.

Solution. Compute:

$$\begin{aligned} b(\boldsymbol{a}) &= 1^{3/2} + 4^{5/2} = 33 \\ \mathsf{D}\, b(\boldsymbol{a}) &= \begin{bmatrix} \frac{3}{2}w^{1/2} & \frac{5}{2}z^{3/2} \end{bmatrix}\Big|_{(w,z)=(1,4)} = \begin{bmatrix} \frac{3}{2} & 20 \end{bmatrix} \\ \mathsf{H}\, b(\boldsymbol{a}) &= \begin{bmatrix} \frac{3}{4}w^{-1/2} & 0 \\ 0 & \frac{15}{4}z^{1/2} \end{bmatrix}\Bigg|_{(w,z)=(1,4)} = \begin{bmatrix} \frac{3}{4} & 0 \\ 0 & \frac{15}{2} \end{bmatrix} \end{aligned}$$

The second order Taylor approximation of $b(w,z)$ at $\boldsymbol{a} = (1,4)$ is:

$$33 + \begin{bmatrix} \frac{3}{2} & 20 \end{bmatrix} \begin{bmatrix} w-1 \\ z-4 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} w-1 & z-4 \end{bmatrix} \begin{bmatrix} \frac{3}{4} & 0 \\ 0 & \frac{15}{2} \end{bmatrix} \begin{bmatrix} w-1 \\ z-4 \end{bmatrix}$$

Specialize the second order Taylor approximation above expression to $(w,z) = (1.02, 3.96)$ to get the approximation

$$33 + \tfrac{3}{2}(0.02) + 20(-0.04) + \tfrac{3}{8}(0.02)^2 + \tfrac{15}{4}(-0.04)^2 = 32.2362$$

for $b(1.02, 3.96)$.

■