M.Eng. 2.6 Mathematics: Directional Derivatives

This sheet can be found on the Web: http://www.ma.ic.ac.uk/~ajm8/MEng26

A function f which associates a scalar value $f(\underline{\mathbf{r}}) \equiv f(x, y, z)$ with every position vector $\underline{\mathbf{r}} \equiv x\,\underline{\mathbf{i}} + y\,\underline{\mathbf{j}} + z\,\underline{\mathbf{k}}$ is said to be a **SCALAR FIELD**, e.g. pressure, temperature.

$\dfrac{\partial f}{\partial x}(\underline{\mathbf{r}})$, $\dfrac{\partial f}{\partial y}(\underline{\mathbf{r}})$ and $\dfrac{\partial f}{\partial z}(\underline{\mathbf{r}})$ are the rates of change of f at the point $\underline{\mathbf{r}}$ in the directions $\underline{\mathbf{i}}$, $\underline{\mathbf{j}}$ and $\underline{\mathbf{k}}$, respectively.

One may be interested in the rate of change of f at the point $\underline{\mathbf{r}}$ in some other direction. The rate of change of f at $\underline{\mathbf{r}}$ in the direction of the unit vector $\underline{\widehat{\mathbf{a}}}$ is said to be the **DIRECTIONAL DERIVATIVE** of f at $\underline{\mathbf{r}}$ in the direction $\underline{\widehat{\mathbf{a}}}$. This generalises the idea of a partial derivative.

CALCULATION of a DIRECTIONAL DERIVATIVE

Let the point P have position vector $\underline{\mathbf{p}} \equiv p_1\underline{\mathbf{i}} + p_2\underline{\mathbf{j}} + p_3\underline{\mathbf{k}}$. and let $\underline{\widehat{\mathbf{a}}} \equiv \widehat{\mathrm{a}}_1\underline{\mathbf{i}} + \widehat{\mathrm{a}}_2\underline{\mathbf{j}} + \widehat{\mathrm{a}}_3\underline{\mathbf{k}}$. Then a point a distance s from P in the direction of $\underline{\widehat{\mathbf{a}}}$ has position vector

$$\underline{\mathbf{r}} \equiv x(s)\underline{\mathbf{i}} + y(s)\underline{\mathbf{j}} + z(s)\underline{\mathbf{k}} = \underline{\mathbf{p}} + s\underline{\widehat{\mathbf{a}}} = (p_1 + s\widehat{\mathrm{a}}_1)\underline{\mathbf{i}} + (p_2 + s\widehat{\mathrm{a}}_2)\underline{\mathbf{j}} + (p_3 + s\widehat{\mathrm{a}}_3)\underline{\mathbf{k}}$$

The derivative of f in the direction of $\underline{\widehat{\mathbf{a}}}$ is therefore

$$\begin{aligned} \lim_{s\to 0} \frac{f(\underline{\mathbf{r}}) - f(\underline{\mathbf{p}})}{|\underline{\mathbf{r}} - \underline{\mathbf{p}}|} &= \lim_{s\to 0} \frac{f(x(s), y(s), z(s)) - f(x(0), y(0), z(0))}{s} \\ &= \lim_{s\to 0} \frac{F(s) - F(0)}{s} = \frac{dF}{ds}(0)\,, \qquad (1) \end{aligned}$$

where $F(s) \equiv f(x(s), y(s), z(s))$. From the CHAIN RULE we have that

$$\begin{aligned} \frac{dF}{ds}(s) &= \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} + \frac{\partial f}{\partial z}\frac{dz}{ds} \\ &= \widehat{\mathrm{a}}_1\frac{\partial f}{\partial x} + \widehat{\mathrm{a}}_2\frac{\partial f}{\partial y} + \widehat{\mathrm{a}}_3\frac{\partial f}{\partial z}\,. \qquad (2) \end{aligned}$$

Introducing

$$\mathrm{grad}\, f(\underline{\mathbf{r}}) \equiv \underline{\nabla}\, f(\underline{\mathbf{r}}) \equiv \frac{\partial f}{\partial x}(\underline{\mathbf{r}})\ \underline{\mathbf{i}} + \frac{\partial f}{\partial y}(\underline{\mathbf{r}})\ \underline{\mathbf{j}} + \frac{\partial f}{\partial z}(\underline{\mathbf{r}})\ \underline{\mathbf{k}}\,; \qquad (3)$$

then from (1) and (2) it follows that

$$\begin{aligned} &\text{the directional derivative of } f \text{ at } \underline{\mathbf{p}} \text{ in the direction } \underline{\widehat{\mathbf{a}}} &= \underline{\widehat{\mathbf{a}}} \cdot \underline{\nabla}\, f(\underline{\mathbf{p}})\,, \\ \Longrightarrow\quad &\text{the directional derivative of } f \text{ at } \underline{\mathbf{r}} \text{ in the direction } \underline{\widehat{\mathbf{a}}} &= \underline{\widehat{\mathbf{a}}} \cdot \underline{\nabla}\, f(\underline{\mathbf{r}})\,, \qquad (4) \end{aligned}$$

where $\underline{\mathbf{r}}$ now denotes a general point. Note that (4) yields that

$$\underline{\mathbf{i}} \cdot \underline{\nabla}\, f(\underline{\mathbf{r}}) \equiv \frac{\partial f}{\partial x}(\underline{\mathbf{r}})\,, \quad \underline{\mathbf{j}} \cdot \underline{\nabla}\, f(\underline{\mathbf{r}}) \equiv \frac{\partial f}{\partial y}(\underline{\mathbf{r}})\,, \quad \underline{\mathbf{k}} \cdot \underline{\nabla}\, f(\underline{\mathbf{r}}) \equiv \frac{\partial f}{\partial z}(\underline{\mathbf{r}})\,. \qquad (5)$$

$\underline{\nabla} \equiv \underline{\mathbf{i}}\,\dfrac{\partial}{\partial x} + \underline{\mathbf{j}}\,\dfrac{\partial}{\partial y} + \underline{\mathbf{k}}\,\dfrac{\partial}{\partial z}$ is a **VECTOR OPERATOR**, $\qquad \underline{\nabla}$: scalar $\to$ vector.

Now

$$\underline{\widehat{\mathbf{a}}} \cdot \underline{\nabla} f(\underline{\mathbf{r}}) = |\underline{\widehat{\mathbf{a}}}|\, |\underline{\nabla} f| \cos\theta = |\underline{\nabla} f| \cos\theta\,, \tag{6}$$

where θ is the angle between $\underline{\widehat{\mathbf{a}}}$ and $\underline{\nabla} f(\underline{\mathbf{r}})$. Therefore choosing $\underline{\widehat{\mathbf{a}}}$ such that $\theta = 0$ yields that

$$\text{the direction of most rapid change for } f \text{ at the point } \underline{\mathbf{r}} \;=\; \frac{\underline{\nabla} f(\underline{\mathbf{r}})}{|\underline{\nabla} f(\underline{\mathbf{r}})|}\,, \tag{7}$$

$$\text{with rate of change} \;=\; |\underline{\nabla} f(\underline{\mathbf{r}})|\,. \tag{8}$$

Example: Let $f(x,y,z) = 3x^2 + xy - z$. What is the rate of change of f in the direction $\underline{\mathbf{i}} + 2\underline{\mathbf{j}} + 3\underline{\mathbf{k}}$ at the point $(1,1,4)$? In what direction is f changing most rapidly in at $(1,1,4)$?

(4), (7) and (8) yield that

$$\underline{\nabla} f(x,y,z) = (6x+y)\,\underline{\mathbf{i}} + x\,\underline{\mathbf{j}} - \underline{\mathbf{k}} \Longrightarrow \underline{\nabla} f(1,1,4) = 7\underline{\mathbf{i}} + \underline{\mathbf{j}} - \underline{\mathbf{k}} \tag{9}$$

$$\text{unit vector in the direction } \underline{\mathbf{i}} + 2\underline{\mathbf{j}} + 3\underline{\mathbf{k}} \text{ is } \quad \underline{\widehat{\mathbf{a}}} = \tfrac{1}{\sqrt{14}}(\underline{\mathbf{i}} + 2\underline{\mathbf{j}} + 3\underline{\mathbf{k}}) \qquad \Longrightarrow$$

$$\text{the rate of change of } f \text{ in the direction of } \underline{\widehat{\mathbf{a}}} \text{ at } (1,1,4) \;=\; \underline{\widehat{\mathbf{a}}} \cdot \underline{\nabla} f(1,1,4) = \tfrac{6}{\sqrt{14}} \approx 1.6\,,$$

$$\text{the direction of most rapid change for } f \text{ at } (1,1,4) \;=\; \frac{\underline{\nabla} f(1,1,4)}{|\underline{\nabla} f(1,1,4)|} = \tfrac{1}{\sqrt{51}}(7\underline{\mathbf{i}} + \underline{\mathbf{j}} - \underline{\mathbf{k}})$$

$$\text{with rate of change} \;=\; |\underline{\nabla} f(1,1,4)| = \sqrt{51} \approx 7.14\,. \tag{10}$$

The curves $f(x,y) =$ constant are called level curves (contours) of f. The surfaces $f(x,y,z) =$ constant are called level surfaces of f, e.g. isotherms, isobars, equipotentials etc.

Let C be a level curve of $f(x,y)$, parameterised by $\underline{\mathbf{r}}(t) = x(t)\underline{\mathbf{i}} + y(t)\underline{\mathbf{j}}$; that is,

$$F(t) \equiv f(x(t), y(t)) = \text{ constant} \Longrightarrow 0 = \frac{dF}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \Longrightarrow \underline{\nabla} f(\underline{\mathbf{r}}(t)) \cdot \frac{d\underline{\mathbf{r}}}{dt}\,. \tag{11}$$

Recall $\dfrac{d\underline{\mathbf{r}}}{dt}(t) = \dfrac{dx}{dt}(t)\underline{\mathbf{i}} + \dfrac{dy}{dt}(t)\underline{\mathbf{j}}$ is tangential to C at $\underline{\mathbf{r}}(t)$. Therefore (11) yields that $\underline{\nabla} f(\underline{\mathbf{r}}(t))$ is perpendicular (normal) to C at $\underline{\mathbf{r}}(t)$. Similarly, in three space dimensions $\underline{\nabla} f(\underline{\mathbf{r}})$ is perpendicular (normal) to the level surface passing through $\underline{\mathbf{r}}$.

Example: Let $f(x,y,z) = 3x^2 + xy - z$. Find the equation of the tangent plane to the level surface $f(x,y,z) = 0$ at (1,1,4).

The above and (9) yield that $\underline{\mathbf{n}} = \underline{\nabla} f(1,1,4) = 7\underline{\mathbf{i}} + \underline{\mathbf{j}} - \underline{\mathbf{k}}$ is the normal to the tangent plane. As $\underline{\mathbf{p}} = \underline{\mathbf{i}} + \underline{\mathbf{j}} + 4\underline{\mathbf{k}}$ is the position vector of the known point on the plane, a general point on the plane, $\underline{\mathbf{r}} = x\underline{\mathbf{i}} + y\underline{\mathbf{j}} + z\underline{\mathbf{k}}$, is such that $(\underline{\mathbf{r}} - \underline{\mathbf{p}}) \,.\, \underline{\mathbf{n}} = 0 \Longrightarrow 7x + y - z = 4$.

In practice one often meets scalar fields that depend only on $r = |\underline{\mathbf{r}}| = \sqrt{x^2 + y^2 + z^2}$, the distance from the origin. Then

$$\underline{\nabla} f(r) = \frac{df}{dr}(r)\left(\frac{\partial r}{\partial x}\underline{\mathbf{i}} + \frac{\partial r}{\partial y}\underline{\mathbf{j}} + \frac{\partial r}{\partial z}\underline{\mathbf{k}}\right) = \frac{1}{r}\frac{df}{dr}(r)(x\underline{\mathbf{i}} + y\underline{\mathbf{j}} + z\underline{\mathbf{k}}) = \frac{1}{r}\frac{df}{dr}(r)\,\underline{\mathbf{r}} \quad , \tag{12}$$

since

$$\frac{\partial}{\partial x}(r^2) = \frac{\partial}{\partial x}(x^2 + y^2 + z^2) \Longrightarrow 2r\frac{\partial r}{\partial x} = 2x \Longrightarrow \frac{\partial r}{\partial x} = \frac{x}{r} \qquad \text{etc.}$$

For example if $f(r) \equiv \ln r$, then (12) yields that $\underline{\nabla} f(r) = \dfrac{1}{r^2}\underline{\mathbf{r}}$.