10 Partial differentiation

(1) Verify that $x^2 = y^2 \sin(yz)$ satisfies $(\partial x/\partial y)_z(\partial y/\partial z)_x(\partial z/\partial x)_y = -1$.

$$x^2 = y^2 \sin(yz) \quad \text{— (1)}$$

$$\frac{\partial}{\partial y_z}(1) \;\Rightarrow\; 2x\left(\frac{\partial x}{\partial y}\right)_z = y^2\cos(yz)\frac{\partial}{\partial y_z}(yz) + 2y\sin(yz)$$

$$\therefore \left(\frac{\partial x}{\partial y}\right)_z = \frac{y^2 z\cos(yz) + 2y\sin(yz)}{2x} \quad \text{— (2)}$$

$$\frac{\partial}{\partial z_x}(1) \;\Rightarrow\; 0 = y^2\cos(yz)\frac{\partial}{\partial z_x}(yz) + 2y\left(\frac{\partial y}{\partial z}\right)_x\sin(yz)$$

$$= y^2\cos(yz)\left[y + z\left(\frac{\partial y}{\partial z}\right)_x\right] + 2y\left(\frac{\partial y}{\partial z}\right)_x\sin(yz)$$

$$\therefore \left(\frac{\partial y}{\partial z}\right)_x = \frac{-y^3\cos(yz)}{y^2 z\cos(yz) + 2y\sin(yz)} \quad \text{— (3)}$$

$$\frac{\partial}{\partial x_y}(1) \;\Rightarrow\; 2x = y^2\cos(yz)\frac{\partial}{\partial x_y}(yz) = y^3\cos(yz)\left(\frac{\partial z}{\partial x}\right)_y$$

$$\therefore \left(\frac{\partial z}{\partial x}\right)_y = \frac{2x}{y^3\cos(yz)} \quad \text{— (4)}$$

$$\therefore (2)\times(3)\times(4) \;\Rightarrow\; \underline{\left(\frac{\partial x}{\partial y}\right)_z\left(\frac{\partial y}{\partial z}\right)_x\left(\frac{\partial z}{\partial x}\right)_y = -1}$$

(2) Consider the function $f(x,y) = x\,y\,(1-y+x)$. Calculate the gradient vector, $\boldsymbol{\nabla} f$, at the points $\left(-\frac{1}{2}, 0\right)$, $\left(-\frac{1}{2}, \frac{1}{2}\right)$ and $\left(0, \frac{1}{2}\right)$. Find the stationary point which lies within the triangle bounded by the points $(-1,0)$, $(0,0)$ and $(0,1)$. Sketch the function within this triangle, and mark in the directions of $\boldsymbol{\nabla} f$ where it has been calculated.

$$\boldsymbol{\nabla} f = \left(\left(\frac{\partial f}{\partial x}\right)_y , \left(\frac{\partial f}{\partial y}\right)_x \right) = \big(y\,(1-y+2x)\ ,\ x\,(1-2y+x) \big)$$

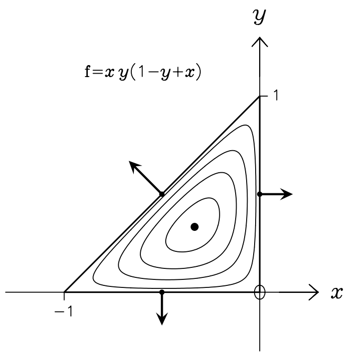
$$\therefore\quad \boldsymbol{\nabla} f\left(-\tfrac{1}{2},0\right) = \underline{\left(0,-\tfrac{1}{4}\right)}\,;\quad \boldsymbol{\nabla} f\left(-\tfrac{1}{2},\tfrac{1}{2}\right) = \underline{\left(-\tfrac{1}{4},\tfrac{1}{4}\right)}\,;\quad \boldsymbol{\nabla} f\left(0,\tfrac{1}{2}\right) = \underline{\left(\tfrac{1}{4},0\right)}$$

For stationary point, $\boldsymbol{\nabla} f = 0$

$$\left.\begin{aligned} \therefore\quad y\,(1-y+2x) &= 0 \\ \text{and}\quad x\,(1-2y+x) &= 0 \end{aligned}\right\} \qquad \begin{aligned} &-1-3x = 0 \\ \therefore\quad &x = -\tfrac{1}{3} \quad\text{and}\quad y = \tfrac{1}{3} \end{aligned}$$

$\therefore$ Stationary point inside triangle is at $\underline{\left(-\frac{1}{3}, \frac{1}{3}\right)}$.

The nature of the stationary point can be ascertained as being a minimum without examining the second derivatives: for example, the gradient vectors point outwards everywhere; or, the value of f at $\left(-\frac{1}{3}, \frac{1}{3}\right)$ is $-\frac{1}{27}$ while the edge of the triangle is a contour with $f=0$.



(3) By using the multivariate form of the Taylor series, derive a version of the Newton-Raphson algorithm for numerically finding a stationary point of a multiparameter function $f(\mathbf{x})$.

Using 'matrix-vector' notation, the Taylor series can be generalized as

$$f(\mathbf{x}) = f(\mathbf{x}_0) + (\mathbf{x}-\mathbf{x}_0)^{T}\,\nabla f(\mathbf{x}_0) + \tfrac{1}{2}(\mathbf{x}-\mathbf{x}_0)^{T}\,\nabla\nabla f(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0) + \cdots$$

where $\nabla\nabla f$ is the matrix of second partial derivatives. For stationary points, $\nabla f(\mathbf{x}) = 0$. Therefore

$$\nabla f(\mathbf{x}) = \nabla f(\mathbf{x}_0) + \nabla\nabla f(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0) + \cdots = 0$$

This is, in fact, just the Taylor series expansion of ∇f. If $\mathbf{x}_0$ is a 'good' estimate of the solution of $\nabla f(\mathbf{x}) = 0$, so that higher-order terms are negligible, then

$$\nabla f(\mathbf{x}_0) + \nabla\nabla f(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0) \approx 0$$

$$\therefore\quad \left[\nabla\nabla f(\mathbf{x}_0)\right]^{-1}\left[\nabla f(\mathbf{x}_0) + \nabla\nabla f(\mathbf{x}_0)\,(\mathbf{x}-\mathbf{x}_0)\right] \approx 0$$

$$\therefore\quad \left[\nabla\nabla f(\mathbf{x}_0)\right]^{-1}\nabla f(\mathbf{x}_0) + \mathbb{I}\,(\mathbf{x}-\mathbf{x}_0) \approx 0$$

$$\therefore\quad \left[\nabla\nabla f(\mathbf{x}_0)\right]^{-1}\nabla f(\mathbf{x}_0) + \mathbf{x} - \mathbf{x}_0 \approx 0$$

$$\text{i.e.}\quad \mathbf{x} \approx \mathbf{x}_0 - \left[\nabla\nabla f(\mathbf{x}_0)\right]^{-1}\nabla f(\mathbf{x}_0)$$

If $f = f(x, y)$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$$

$$\nabla\nabla f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\,\partial y} \\ \frac{\partial^2 f}{\partial y\,\partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

In other words, a better estimate of the solution to $\nabla f(\mathbf{x}) = 0$ can be obtained from the values of the gradient vector and the Hessian matrix at $\mathbf{x}_0$. This forms the basis of an iterative Newton-Raphson algorithm:

$$\underline{\mathbf{x}_{N+1} = \mathbf{x}_N - \left[\nabla\nabla f(\mathbf{x}_N)\right]^{-1}\nabla f(\mathbf{x}_N)}$$

where $\mathbf{x}_N$ is the estimate of the solution at the N^{th} iteration, and $\mathbf{x}_{N+1}$ is the improved update.

The Newton-Raphson algorithm above is usually by far the most efficient way of numerically homing in on a stationary point given a good initial guess. The last proviso is important, though, because the procedure can easily diverge rapidly away from the solution being sought if the initial estimate is not close-enough to the desired point.

If $f = f(x_1, x_2, x_3, \cdots, x_M) = f(\mathbf{x})$

$$\left[\nabla f(\mathbf{x}_0)\right]_i = \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}_0}$$

$$\left[\nabla\nabla f(\mathbf{x}_0)\right]_{ij} = \left.\frac{\partial^2 f}{\partial x_i\,\partial x_j}\right|_{\mathbf{x}_0}$$