1 Partial and directional derivatives

1.1 Partial derivatives

The derivative of a function with one input is a single value. The derivative of a function of two variables, however, can have multiple values depending on what you differentiate with respect to - think of this as finding the gradient along different directions of a landscape. This can then be generalised to functions with n inputs, taking the derivative with respect to any vector, which will give the rate of change of the function in that direction. This process of finding a derivative of a multivariate function (a function with multiple inputs) is called partial differentiation. The partial derivative in the x direction is denoted by $\frac{\partial f}{\partial x}$ or $f_x(x, y)$.

To find the partial derivative of a function with respect to one of its variables, simply treat the corresponding variable as the variable to be differentiated, and all other variables as constants. Differentiating a function $f(x, y)$ with respect to x involves taking small steps along the corresponding graph of the function in the x direction and working out an expression for the rate of change of the function in that direction. If we are differentiating with respect to x, we only move in the x direction and there is no change whatsoever in the y direction, meaning that y can be treated as a constant. This is akin to taking a slice through the function parallel to the x axis and treating this as a function of one input.

Consider the function $f(x, y) = 5x^3 + 2xy^2 + 3y$. To calculate $\frac{\partial f}{\partial x}$, treat x as the variable and y as a constant; $\frac{\partial f}{\partial x} = 15x^2 + 2y^2$. $\frac{\partial f}{\partial y}$ is calculated in a similar way, with y as the variable and x as a constant; $\frac{\partial f}{\partial y} = 4xy + 3$.

1.2 Gradient

For a multivariate function, the gradient is a vector in the direction of the largest rate of change (the "steepest" direction), whose length gives the rate of change. The gradient of a function f is denoted ∇f.

The gradient of $f(x, y, z)$ is defined as $\nabla f = \frac{\partial f}{\partial x}\hat{\mathbf{x}} + \frac{\partial f}{\partial y}\hat{\mathbf{y}} + \frac{\partial f}{\partial z}\hat{\mathbf{z}}$, where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are the unit vectors in the x, y and z directions. This essentially produces a gradient vector in each direction (x, y and z) and produces a resultant vector of these three gradients, giving the overall gradient of the function.

1.3 Directional derivatives

To calculate the directional derivative of a function, find the partial derivative with respect to each basis vector and multiply by the corresponding component of the unit vector in the direction you are interested in, then sum these. This is equivalent to finding the dot product of the gradient and the unit vector in the direction in which you are interested. For a 2-input function:

$$\begin{aligned}
\nabla f &= \frac{\partial f}{\partial x}\hat{\mathbf{x}} + \frac{\partial f}{\partial y}\hat{\mathbf{y}} \\
\hat{\mathbf{u}} &= (u_x\hat{\mathbf{x}} + u_y\hat{\mathbf{y}}) \\
\therefore \nabla f \cdot \hat{\mathbf{u}} &= \frac{\partial f}{\partial x}u_x + \frac{\partial f}{\partial y}u_y \\
&= \frac{\partial f}{\partial u}
\end{aligned}$$

Consider our previous example of $f(x, y) = 5x^3 + 2xy^2 + 3y$, along with the vector $\mathbf{u} = 3\hat{\mathbf{x}} + 4\hat{\mathbf{y}}$. $\hat{\mathbf{u}}$ is the unit vector in the direction of $\mathbf{u}$, $\hat{\mathbf{x}}$ is the unit vector in the direction of x and $\hat{\mathbf{y}}$ is the unit vector in the direction of y.

$$\begin{aligned}
\hat{\mathbf{u}} &= \frac{1}{\sqrt{3^2 + 4^2}}(3\hat{\mathbf{x}} + 4\hat{\mathbf{y}}) \\
&= \frac{1}{5}(3\hat{\mathbf{x}} + 4\hat{\mathbf{y}}) \\
\frac{\partial f}{\partial x} &= 15x^2 + 2y^2, \frac{\partial f}{\partial y} = 4xy + 3 \\
\frac{\partial f}{\partial u} &= \nabla f \cdot \hat{\mathbf{u}} = \frac{\partial f}{\partial x}u_x + \frac{\partial f}{\partial y}u_y \\
&= \frac{1}{5} \times 3 \times (15x^2 + 2y^2) + \frac{1}{5} \times 4 \times (4xy + 3) \\
&= \frac{1}{5}(45x^2 + 6y^2 + 16xy + 12)
\end{aligned}$$

2 Vector fields

A vector field is the assignment of a vector to every point in space. The vector field itself can represent many things - the force of gravity throughout space, the wind speed and direction on a weather map, the flow of fluid around an object, etc., etc. The field itself can be visually represented by evenly spaced arrows throughout space, whose direction and length represent the corresponding properties of the vector at that point. It must be remembered that representing a vector field in this way is only an approximation. To draw an arrow at every point in space would mean they'd all overlap, so they are drawn at regular intervals with the assumption that the vector field between arrows is fairly constant and hence can be deduced from the surrounding arrows.

Another intuition for vector fields which is useful for the following sections is to think of them as representing fluid flow. Although vector fields can be used for other things too, fluid flow is easier to picture than gravitational field lines.

3 Divergence

The divergence of a vector field at a given point is a measure of how much the vectors around that point are directed away from it - how much they diverge. If the vector field is pictured as representing fluid flow; areas of positive divergence are sources of fluid - fluid flows away from them more than it does towards them; areas of negative divergence are sinks - fluid flows towards them more than it does away from them; and areas of zero divergence are where the fluid has no net gain or loss. It is important to note that for an area of positive divergence, vectors can still point towards it; the requirement is that the sum of the outward vectors is greater than the sum of inward vectors.

Divergence is the dot product of the vector field and its gradient, so the divergence of the vector field $\mathbf{F}$ is denoted $\nabla \cdot \mathbf{F}$. When written in full for a three-dimensional vector field, this gives:

$$\begin{aligned}\nabla &= \frac{\partial f}{\partial x}\hat{\mathbf{x}} + \frac{\partial f}{\partial y}\hat{\mathbf{y}} + \frac{\partial f}{\partial z}\hat{\mathbf{z}} \\ \mathbf{F} &= (F_x, F_y, F_z) \\ \therefore \nabla \cdot \mathbf{F} &= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\end{aligned}$$

To see why divergence can be written like this, picture an small cube (or n-dimensional analog) around the point in question, aligned with the coordinate system's axes. The overall "fluid" leaving the box is therefore determined by summing the differences of the vector field's value on opposite sides. As the size of the cube approaches zero, this sum is simply the sum of the partial derivative with respect to each axis - the sum of the rate of change of fluid flow in each direction. The dot product operator produces a scalar result, so the divergence of a vector field is a scalar field.

The dot product, in essence, measures how close two vectors are to being parallel. So, the divergence is a measure of how parallel the gradient and the direction of flow of the vector field are.

4 Curl

The curl of a vector field at a given point is a measure of how much the vectors around that point are directed around it - how much they curl around it. Using the picture of fluid flow again; in areas of positive curl the fluid moves anticlockwise more than clockwise; in areas of negative curl it moves clockwise more than anticlockwise; and in areas of zero curl there is no net rotation around the point. As with divergence, it is important to note that curl represents the net value, so fluid can still rotate clockwise in an area of positive curl - but it must be outweighed by the anticlockwise rotation of fluid. Curl is defined only in three dimensions.

Curl is the cross product of the vector field and its gradient, so the curl of a vector field $\mathbf{F}$ is denoted $\nabla \times \mathbf{F}$. Written in full, this gives:

$$\begin{aligned}\nabla &= \frac{\partial f}{\partial x}\hat{\mathbf{x}} + \frac{\partial f}{\partial y}\hat{\mathbf{y}} + \frac{\partial f}{\partial z}\hat{\mathbf{z}} \\ \mathbf{F} &= (F_x, F_y, F_z) \\ \therefore \nabla \times \mathbf{F} &= \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}\right)\hat{\mathbf{x}} + \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}\right)\hat{\mathbf{y}} + \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}\right)\hat{\mathbf{z}}\end{aligned}$$

The coefficient of each basis vector is the rotation around that vector. Consider an anti-clockwise vector field curling around a point. As the value of x increases, the y component of the vector increases - a positive partial derivative of the F_y with respect to x. In a similar way, as the value of y increases the x component of the vector decreases - a negative partial derivative of F_x with respect to y. Therefore if $\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y}$ is positive, there is anti-clockwise rotation, represented by positive curl values. Considering this in each axis plane, the overall curl vector may be obtained. The direction is the axis of rotation and the magnitude represents the rate of rotation. This is where the right-hand rule stems from.

The cross product operator essentially measures how perpendicular two vectors are, so curl is a measure of how perpendicular the gradient and flow of a vector field are - hence how much they curl.