EE2 Mathematics : Functions of Multiple Variables

http://www2.imperial.ac.uk/∼nsjones

These notes are not identical word-for-word with my lectures which will be given on the blackboard. Some of these notes may contain more examples than the corresponding lecture while in other cases the lecture may contain more detailed working **you are therefore strongly advised to attend lectures**.

Contents

1 Partial Differentiation and Multivariable Functions

In the following we will be considering functions of multiple variables $f(x, y...)$. We will principally consider the functions of just two variables, $f(x, y)$, but most of the concepts discussed can be generalized to multiple variables. In the lectures you will be introduced to perspective plots and (constant) contour plots. This part of the course can be viewed as introducing the basics of optimization: something critical for quantitative design.

1.1 Partial Differentials Reminder

From starting point (x_0, y_0), we will consider small changes in the x, y-plane of size $(\Delta x, \Delta y)$ and the small changes that they correspond to in f, of size Δf.

Consider small changes to x with $y = constant = y_0$. It follows that $\Delta f = f(x_0 + \Delta x, y_0) - f(x_0, y_0)$ and the instantaneous slope in the x-direction is then

$$\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x} = \left.\frac{\partial f}{\partial x}\right|_{y=y_0} \tag{1.1}$$

(if the limit exists). This is called the partial derivative of f with respect to x while keeping y constant. There are a variety of notations for partial derivatives e.g.:

$$\left.\frac{\partial f}{\partial x}\right|_y = \left(\frac{\partial f}{\partial x}\right)_y = \frac{\partial f}{\partial x} = f_x \tag{1.2}$$

note that in the last two equivalences, the fact that y is held constant is simply assumed. The partial differential with respect to y, $\left.\frac{\partial f}{\partial y}\right|_{x=x_0}$ is defined similarly.

Example: If $f(x, y) = x^3 \sin(xy)$ then $\left.\frac{\partial f}{\partial x}\right|_y = 3x^2 \sin(xy) + x^3 y \cos(xy)$ and also $\left.\frac{\partial f}{\partial y}\right|_x = x^4 \cos(xy)$.

We can also take the partial derivative of a partial derivative yielding expressions like $\frac{\partial^2 f}{\partial x^2} = f_{xx}$ or $\frac{\partial^2 f}{\partial x \partial y} = f_{xy}$.

1.2 The Total Differential

What is the change in f for any small step $(\Delta x, \Delta y)$? We'll use a relatively standard argument (in the lectures we'll use a graphical argument) to obtain the total differential:

$$\begin{aligned}
\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\
\Delta f &= f(x + \Delta x, y + \Delta y) + (-f(x, y + \Delta y) + f(x, y + \Delta y)) - f(x, y) \\
\Delta f &= \Delta x \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x} + \Delta y \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}
\end{aligned} \tag{1.3}$$

This resembles the definitions for the partial differentials used previously. For $(\Delta x, \Delta y)$ sufficiently small it becomes

$$\Delta f \approx \frac{\partial f(x,y)}{\partial x}\Delta x + \frac{\partial f(x,y)}{\partial y}\Delta y. \tag{1.4}$$

(the first term might seem to approximate $\frac{\partial f(x,y+\Delta)}{\partial y})\Delta x$ but this in turn approximates $\frac{\partial f(x,y)}{\partial x}\Delta x$). In the limit $\Delta x, \Delta y \to 0$ this becomes the **Total Differential**:

$$\boxed{df = \frac{\partial f(x,y)}{\partial x}dx + \frac{\partial f(x,y)}{\partial y}dy} \tag{1.5}$$

where df, dx, dy are called abstract differentials. This generalises to $df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \ldots \frac{\partial f}{\partial x_n}dx_n$ for functions of n variables.

Example: Consider the function $f(x,y) = x^2 + xy - 3y^2$. The constant contour $f = 1$ passes through $(0,1)$. What is the value of y on the same contour when $x = 0.1$?

Along the contour $\Delta f = 0$. It follows from Eq. 1.4 that $\frac{\partial f(x,y)}{\partial x}\Delta x + \frac{\partial f(x,y)}{\partial y}\Delta y = 0 = (2x+y)\Delta x + (x-6y)\Delta y$ evaluated at $(0,1)$ so $\Delta y \approx \frac{\Delta x}{6}$ and so $y \approx 1 + \Delta y \approx 1.017$.

1.3 The Chain Rule

This considers the setting where $f(x,y)$ but x and y are both themselves functions of a variable u such that $x(u), y(u)$. This is a natural setting e.g. one might have a function (like fuel consumption) depending on distance and speed but, in turn, both of these quantities might be functions of time. E.g. $f = bx + xy$ where $x = au^2/2$ and $y = au$. Since $f(x,y) = f(u)$ it follows that $\frac{df}{du}$ is well defined (since f is a univariate function of u).

$$\frac{df}{du} = \lim_{\Delta u \to 0}\frac{\Delta f}{\Delta u} = \lim_{\Delta u \to 0}\frac{\frac{\partial f(x,y)}{\partial x}\Delta x + \frac{\partial f(x,y)}{\partial y}\Delta y}{\Delta u} \tag{1.6}$$

from this follows the **Chain Rule**:

$$\boxed{\frac{df}{du} = \frac{\partial f(x,y)}{\partial x}\frac{dx}{du} + \frac{\partial f(x,y)}{\partial y}\frac{dy}{du}.} \tag{1.7}$$

Note that this only applies when f is a univariate function of u.

Example: From the example above and using Eq. (1.7) $\frac{df}{du} = (b+y).au + x.a$ which simplifies to $bau + \frac{3}{2}a^2u^2$.

The chain rule generalizes to functions of more variables as $\frac{df}{du} = \frac{\partial f}{\partial x_1}\frac{dx_1}{du} + \frac{\partial f}{\partial x_2}\frac{dx_2}{du} + \ldots .\frac{\partial f}{\partial x_n}\frac{dx_n}{du}$.

2 Change of Variables

Suppose $f = f(x,y)$ and that $x = x(u,v)$ and $y = y(u,v)$ (this is a generalization of the previous situation where we had $x(t)$ and $y(t)$). So we can write $f = f(x,y) = f(x(u,v), y(u,v)) = f(u,v)$. It is useful to be able to toggle between these co-ordinate systems when trying to solve problems. We can write out various versions of the total differential:

$$df = \left.\frac{\partial f}{\partial u}\right|_v du + \left.\frac{\partial f}{\partial v}\right|_u dv \quad (\text{since} \quad f(u,v)) \tag{2.1}$$

$$df = \left.\frac{\partial f}{\partial x}\right|_y dx + \left.\frac{\partial f}{\partial y}\right|_x dy \quad (\text{since} \quad f(x,y)) \tag{2.2}$$

$$dx = \left.\frac{\partial x}{\partial u}\right|_v du + \left.\frac{\partial x}{\partial v}\right|_u dv \quad (\text{since} \quad x(u,v)) \tag{2.3}$$

$$dy = \left.\frac{\partial y}{\partial u}\right|_v du + \left.\frac{\partial y}{\partial v}\right|_u dv \quad (\text{since} \quad y(u,v)) \tag{2.4}$$

We can now substitute (2.3) and (2.4) into (2.2) obtaining:

$$df = \left.\frac{\partial f}{\partial x}\right|_y \left(\left.\frac{\partial x}{\partial u}\right|_v du + \left.\frac{\partial x}{\partial v}\right|_u dv\right) + \left.\frac{\partial f}{\partial y}\right|_x \left(\left.\frac{\partial y}{\partial u}\right|_v du + \left.\frac{\partial y}{\partial v}\right|_u dv\right) \tag{2.5}$$

Grouping terms in (2.5) and comparing terms in du and dv with (2.1) we find:

$$\left.\frac{\partial f}{\partial u}\right|_v = \left.\frac{\partial f}{\partial x}\right|_y \left.\frac{\partial x}{\partial u}\right|_v + \left.\frac{\partial f}{\partial y}\right|_x \left.\frac{\partial y}{\partial u}\right|_v \tag{2.6}$$

which "expresses a change of variables in f". And similarly

$$\left.\frac{\partial f}{\partial v}\right|_u = \left.\frac{\partial f}{\partial x}\right|_y \left.\frac{\partial x}{\partial v}\right|_u + \left.\frac{\partial f}{\partial y}\right|_x \left.\frac{\partial y}{\partial v}\right|_u \tag{2.7}$$

Rule of thumb: One can obtain change of variables expressions from the total differential by "multiplying through by $\left.\frac{}{\partial v}\right|_u$". I.e. we can obtain (2.6) by writing "$\partial f = \left.\frac{\partial f}{\partial x}\right|_y \partial x + \left.\frac{\partial f}{\partial y}\right|_x \partial y$" and "multiplying through by $\left.\frac{}{\partial v}\right|_u$". Note that we're using "∂x" and "∂y" and not dx and dy because $x = x(u,v)$ and $y = y(u,v)$.

Example of a change of variables:

Suppose I want to go from thinking about a function in polar co-ordinates, $f(r,\theta) = r^2 \sin\theta$ to representing it in Cartesian co-ordinates.

Find $\left.\frac{\partial f}{\partial x}\right|_y$ where $x = r\cos\theta$, $y = r\sin\theta$, $r = \sqrt{(x^2+y^2)}$ and $\theta = \tan^{-1} y/x$. Writing the total differential for $f(\theta, r)$ as $df = \left.\frac{\partial f}{\partial r}\right|_\theta dr + \left.\frac{\partial f}{\partial \theta}\right|_r d\theta$. Using the rule of thumb above we have

$$\left.\frac{\partial f}{\partial x}\right|_y = \left.\frac{\partial f}{\partial r}\right|_\theta \left.\frac{\partial r}{\partial x}\right|_y + \left.\frac{\partial f}{\partial \theta}\right|_r \left.\frac{\partial \theta}{\partial x}\right|_y \tag{2.8}$$

it follows that

$$\left.\frac{\partial f}{\partial x}\right|_y = 2r\sin\theta \cdot \frac{1}{2}(x^2+y^2)^{-\frac{1}{2}} \cdot 2x + r^2\cos\theta \cdot \frac{1}{1+y^2/x^2} . \frac{-y}{x^2} \tag{2.9}$$

This simplifies to $\left.\frac{\partial f}{\partial x}\right|_y = xy/\sqrt{x^2+y^2}$.

Note[1] What about double partial differentials? $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(\frac{\partial f}{\partial x}) = [\frac{\partial}{\partial x}][\frac{\partial}{\partial x}]f$. So

$$\frac{\partial^2 f}{\partial x^2} = \left[\frac{x}{r}\left.\frac{\partial}{\partial r}\right|_\theta + \frac{-y}{r^2}\left.\frac{\partial}{\partial \theta}\right|_r\right]\left[\frac{x}{r}\left.\frac{\partial}{\partial r}\right|_\theta + \frac{-y}{r^2}\left.\frac{\partial}{\partial \theta}\right|_r\right] f \tag{2.10}$$

.

How should I interpret expressions like this? E.g. what is the first term when I expand the above?

$$\begin{aligned} \frac{x}{r}\frac{\partial}{\partial r}\left[\frac{x}{r}\frac{\partial}{\partial r}\right] f &= \frac{x}{r}\frac{\partial}{\partial r}\left[\frac{x}{r}\frac{\partial f}{\partial r}\right] && (2.11) \\ &= \frac{x}{r}\left[\frac{\partial(x/r)}{\partial r}\frac{\partial f}{\partial r} + \frac{x}{r}\frac{\partial^2 f}{\partial r^2}\right] && (2.12) \end{aligned}$$

How do we obtain equations like (2.10)? Well let's consider Equation (2.8) and $\frac{\partial^2 f}{\partial x^2}$. Writing $g(r,\theta) = \left.\frac{\partial f}{\partial x}\right|_y$ we are then at liberty to consider the total differential of $dg(r,\theta)$ in terms of the abstract differentials dr and $d\theta$. By construction, $\frac{\partial^2 f}{\partial x^2} = \frac{\partial g}{\partial x}$ and we can obtain $\frac{\partial g}{\partial x}$ by using the above rule of thumb and "multiplying $dg(r,\theta)$ through by $\frac{\partial}{\partial x}$": this yields Eq. (2.8) in terms of g not f. Now we can take this expression for $\frac{\partial g}{\partial x}$ and eliminate g using $g(r,\theta) = \left.\frac{\partial f}{\partial x}\right|_y$. One then obtains an expression similar to Equation (2.10).

Note[2] Implicit expressions have cropped up already e.g. $f(x,y,z) = 0$. First of all it might be useful to remember how, in the univariate case you have been taught to do implicit differentation. E.g. Remind yourself how you would find $\frac{dy}{dx}$ when $y^2x + x\log y + 3 = 0$ (hint: you don't do it by rearranging for an expression $y = h(x)$). So what about the multivariate case if I want to find e.g. $\left.\frac{\partial z}{\partial x}\right|_y$? Example: for $y = z^3 + xz$. Differentiate with respect to x holding y constant (and noting the obvious, but perhaps helpful to some, that $[\left.\frac{\partial}{\partial x}\right|_y]z = \left.\frac{\partial z}{\partial x}\right|_y$). It follows from the expression for y that $0 = 3z^2\left.\frac{\partial z}{\partial x}\right|_y + \left.\frac{\partial x}{\partial x}\right|_y z + x\left.\frac{\partial z}{\partial x}\right|_y$. We can then rearrange and solve for $\left.\frac{\partial z}{\partial x}\right|_y = \frac{-z}{3z^2+x}$.

Note[3] There are two identities which hold for $f(x,y,z) = 0$ which you may have encountered previously. These are called the reciprocity relation and the cyclic relation.

$$\begin{aligned} \left.\frac{\partial x}{\partial y}\right|_z &= \left[\left.\frac{\partial y}{\partial x}\right|_z\right]^{-1} && (2.13) \\ \frac{\partial x}{\partial y}|_z . \frac{\partial y}{\partial z}|_x . \frac{\partial z}{\partial x}|_y &= -1 && (2.14) \end{aligned}$$

Try and obtain these for yourself. This can be done by writing out the total differentials for $dx(y,z)$ and $dy(x,z)$ and using the expression for dy to eliminate dy from the expression for

dx. Then, by considering the scenarios first when z is constant and then when x is constant you can obtain the above.

3 Taylor's theorem for multi-variable functions

Reminder: in univariate case

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2!}f''(x_0) + ... + \frac{(x - x_0)^{n-1}}{(n-1)!}f^{n-1}(x_0) + ... \quad (3.1)$$

where we could also have written $x - x_0 = \Delta x$.

Generalizing to two variables:

$$f(x, y) = f(x_0, y_0) + \Delta x\frac{\partial f}{\partial x} + \Delta y\frac{\partial f}{\partial y} + \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x\Delta y + \frac{\partial^2 f}{\partial y^2}(\Delta y)^2\right] + ... \quad (3.2)$$

Where we define $\Delta x = x - x_0$ and $\Delta y = y - y_0$ and evaluate the derivatives at x_0, y_0. We could write the quadratic term in different notation as $(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y})(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y})f$ where $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ act only on f not on Δx or Δy. The above expansion then becomes:

$$f(x, y) = \sum_{n=0}^{\infty}\frac{1}{n!}\left[\left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^n f(x, y)\right]_{x_0, y_0} \quad (3.3)$$

For more than two dimensions we can write this as

$$f(\mathbf{x}) = \sum_{n=0}^{\infty}\frac{1}{n!}\left[(\Delta\mathbf{x}\cdot\nabla)^n f(\mathbf{x})\right]_{\mathbf{x}=\mathbf{x_0}} \quad (3.4)$$

or

$$f(\mathbf{x}) = f(\mathbf{x_0}) + \sum_i \frac{\partial f}{\partial x_i}\Delta x_i + \frac{1}{2!}\sum_i\sum_j \frac{\partial^2 f}{\partial x_i \partial x_j}\Delta x_i\Delta x_j + \dots \quad (3.5)$$

Hessian: If $f(x_1...x_n)$ we can define $H_{ij} = \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}f(x_1, ..., x_n)$. The matrix H is called the Hessian. H_{ij} appears in the equation above (3.5).

4 Gradient

The gradient vector is called "grad f" or $\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$. If $f = f(u_1...u_n)$ then the same notation holds for higher dimensions and the i^{th} component of $\nabla f_i = \frac{\partial f}{\partial u_i}$.

The total differential $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ can be written as $df = d\underline{s}\cdot\nabla f$ where $d\underline{s} = (dx, dy)$. Instead of $d\underline{s}$ we could also think about small (finite) steps $\Delta\underline{S} = (\Delta x, \Delta y)$ or $|\Delta\underline{S}|\widehat{\Delta\underline{S}}$ (where $\widehat{\Delta\underline{S}}$ is a unit vector in direction $\Delta\underline{S}$). We can thus write

$$\Delta f \simeq \Delta \underline{S} \cdot \nabla f = |\Delta \underline{S}| \widehat{\Delta \underline{S}} \cdot \nabla f \tag{4.1}$$

Rewriting the dot product:

$$\Delta f \simeq |\nabla f||\Delta \underline{S}| \cos\theta \tag{4.2}$$

where θ is the angle between vectors ∇f and $\Delta \underline{S}$. Equation 4.2 is useful: it relates the change in the function f, Δf, to ∇f and $\Delta \underline{S}$. If $\Delta \underline{S}$ is a step which is parallel to ∇f then θ is zero and Δf is maximized. This means that small steps in a direction parallel to ∇f are in the direction of greatest change of f. This allows an interpretation of ∇f: it points in the direction of greatest change of f. Similarly if $\Delta \underline{S}$ is perpendicular to ∇f then Equation 4.2 tells us that the change in f is zero. In this case $\Delta \underline{S}$ is a motion along a constant contour of f (or we're located at one of f's stationary points - see later).

This gives us an interpretation for the *direction* of ∇f but what about its *magnitude*? $|\nabla f(x,y)|$ tells us the gradient in the direction $\nabla f(x,y)$. In Eq. 4.2 if $\Delta \underline{S} \parallel \nabla f$ then $\theta = 0$ and so rearranging we find that $|\nabla f| \simeq \frac{\Delta f}{|\Delta \underline{S}|}$. In the limit of infinitesimal steps $d\underline{s}$ in the direction ∇f this becomes $\frac{df}{d\underline{s}} = |\nabla f|$. I.e. previously we took slices through our landscape which were parallel to the co-ordinate axes, e.g. setting $y = y_0$. We could then calculate gradients like $\frac{\partial f}{\partial x}|_{y=y_0}$. But we can also, for any point (x,y), take a slice through the function in a direction $\nabla f(x,y)$ which passes through (x,y). The instantaneous rate of change of f along this slice, and at (x,y), is the value of $|\nabla f(x,y)|$.

Recap: $\nabla f(x,y)$ is a vector pointing in the direction of greatest change of f at the point (x,y). Its magnitude is equal to the rate of change of f in this direction (at (x,y)).

From above, if $d\underline{s}$ is not parallel to f then a step, $d\underline{s}$, of infinitesimal size in direction $\widehat{\underline{a}}$ has a rate of change of f specified by $\frac{df}{d\underline{s}} = \nabla f \cdot \widehat{\underline{a}} = |\nabla f| \cos\theta$ where θ is the angle between $\widehat{\underline{a}}$ and ∇f.

A standard example: Consider the function $f = x^2 y + yz$. Find i) ∇f ii) Find the rate of change of f ($\frac{df}{d\underline{s}}$) in a direction **a** $= (1,2,3)$ at point $(1,2,-1)$ iii) what is the direction of the largest rate of change at this point, and what is its magnitude?

i) $\nabla f = (2xy, x^2 + z, y)$. ii) First find $\widehat{\underline{a}}$ then $\frac{df}{d\underline{s}} = \nabla f \cdot \widehat{\underline{a}}$. At $(1,2,-1)$ $\nabla f = (4,0,2)$ so $\frac{df}{d\underline{s}} = \frac{10}{\sqrt{14}}$. iii) This is the direction $\nabla f = (4,0,2)$ and of magnitude $|\nabla f| = \sqrt{(}20)$.

5 Stationary Points

Univariate case: When we differentiate $f(x)$ and solve for the x_0 such that $\frac{df}{dx} = 0$, the gradient is zero at these points. What about the rate of change of gradient: $\frac{d}{dx}(\frac{df}{dx})$ at the minimum x_0? For a minimum the gradient increases as $x_0 \to x_0 + \Delta x$ ($\Delta x > 0$). It follows that $\frac{d^2f}{dx^2} > 0$. The opposite is true for a maximum: $\frac{d^2f}{dx^2} < 0$, the gradient decreases upon positive steps away from x_0. For a point of inflection $\frac{d^2f}{dx^2} = 0$.

Multivariate case: Stationary points occur when $\nabla f = \underline{0}$. In 2-d this is $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}) = \underline{0}$, namely, a generalization of the univariate case. Recall that $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$ can be written as $df = d\underline{s} \cdot \nabla f$ where $d\underline{s} = (dx, dy)$. If $\nabla f = \underline{0}$ at (x_0, y_0) then *any* infinitesimal step $d\underline{s}$ away from (x_0, y_0) will still leave f unchanged, i.e. $df = 0$.

There are three types of stationary points of $f(x, y)$: Maxima, Minima and Saddle Points. We'll draw some of their properties on the board. We will now attempt to find ways of identifying the character of each of the stationary points of $f(x, y)$.

Consider a Taylor expansion about a stationary point (x_0, y_0). We know that $\nabla f = \underline{0}$ at (x_0, y_0) so writing $(\Delta x, \Delta y) = (x - x_0, y - y_0)$ we find:

$$\begin{aligned} \Delta f &= f(x,y) - f(x_0,y_0) \\ &\simeq 0 + \frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(\Delta x)^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}(\Delta y)^2\right]. \end{aligned} \tag{5.1}$$

Maxima: At a maximum *all* small steps away from (x_0, y_0) lead to $\Delta f < 0$. From Eq. (5.1) it follows that for a maximum:

$$\Delta x^2 f_{xx} + 2\Delta x \Delta y f_{xy} + \Delta y^2 f_{yy} < 0 \tag{5.2}$$

for all $(\Delta x, \Delta y)$. Since this holds for all $(\Delta x, \Delta y)$ this includes $\Delta y = 0$. It follows that

$$f_{xx} < 0 \tag{5.3}$$

(also $f_{yy} < 0$ by similar arguments). Since Eq. (5.2) holds for arbitrary $(\Delta x, \Delta y)$, it must also hold for $(\Delta x, \Delta y) = (\lambda \Delta y, \Delta y)$ (i.e. along the locus $\Delta x = \lambda \Delta y$). In this case Eq. (5.2) becomes:

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} < 0. \tag{5.4}$$

Multiplying through by $f_{xx} < 0$ we obtain:

$$\lambda^2 f_{xx}^2 + 2\lambda f_{xx} f_{xy} + f_{xx} f_{yy} > 0 \tag{5.5}$$

$$f_{xy}^2 - f_{xx} f_{yy} < (f_{xx}\lambda + f_{xy})^2, \tag{5.6}$$

for all λ (this last step is a useful trick). The right-hand side of Equation (5.6) can be set to zero (by picking λ appropriately) but can't be smaller (because it is a squared term) so:

$$f_{xy}^2 - f_{xx} f_{yy} < 0. \tag{5.7}$$

Minima: The same type of arguments apply for minima where *all* small steps from (x_0, y_0) lead to $\Delta f > 0$.

$$\begin{aligned} f_{xx} &> 0 & (5.8) \\ f_{xy}^2 - f_{xx}f_{yy} &< 0. & (5.9) \end{aligned}$$

(and also $f_{yy} > 0$).

Saddle Points: In the Taylor expansion for Δf what happens if $\Delta x = \lambda \Delta y$ such that Eq. (5.1) is zero to second order?

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} = 0. \quad (5.10)$$

We can use the same trick of multiplying through by f_{xx} and rearranging as we did in Eqs. (5.5,5.6) and noting that $(f_{xx}\lambda + f_{xy})^2 > 0$ for all λ we find:

$$f_{xy}^2 - f_{xx}f_{yy} > 0 \quad (5.11)$$

which is the condition for a stationary point to be a saddle point. Note that f_{xx}, f_{yy} are unconstrained. Eq. (5.10) can be solved for real roots λ_1, λ_2. This specifies two directions $\Delta x = \lambda_1 \Delta y$ and $\Delta x = \lambda_2 \Delta y$ on which $f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0)$ to second order (these correspond the the asymptotes of the saddle's locally hyperbolic contours).

Summary: The sufficient conditions for a stationary point to be a Max, Min, Saddle are:

$$\begin{aligned} Max: \quad f_{xx} < 0 \quad \text{and} \quad f_{xy}^2 - f_{xx}f_{yy} &< 0 & (5.12) \\ Min: \quad f_{xx} > 0 \quad \text{and} \quad f_{xy}^2 - f_{xx}f_{yy} &< 0 & (5.13) \\ Saddle: \quad f_{xy}^2 - f_{xx}f_{yy} &> 0 & (5.14) \end{aligned}$$

Note that these are not necessary conditions: consider $f(x,y) = x^4 + y^4$ (a question in one of your examples classes). To classify the stationary points in such cases the Taylor expansion used in Eq. (5.1) must be taken to higher order.

A standard example: Find and classify the stationary points of $f(x,y) = x^3 - 3x^2 + 2xy - y^2$ and sketch its contours.

From $\nabla f = \underline{0}$ it follows that $f_x = 3x^2 - 6x + 2y = 0$ and $f_y = 2x - 2y = 0$. It follows that $x = y$ and so that $3x^2 - 4x = 0$ and thus $x = 0, \frac{4}{3}$. Stationary points are thus $(0,0)$ and $(\frac{4}{3}, \frac{4}{3})$.

We can calculate the possible second derivatives (equivalent to finding the Hessian mentioned earlier): $f_{xx} = 6x - 6$; $f_{xy} = 2$; $f_{yy} = -2$. At $(0,0)$ by substituting in the relevant values one finds $f_{xx} < 0$ and $f_{xy}^2 - f_{xx}f_{yy} < 0$. Thus $(0,0)$ is a maximum. At $(\frac{4}{3}, \frac{4}{3})$ we find we have a saddle point, since $f_{xy}^2 - f_{xx}f_{yy} > 0$.

In order to sketch this we need to find the asymptotes of the locally hyperbolic contours about the saddle point. Which two directions have $\Delta f = 0$ (to second order) at the saddle point? I.e. what $\Delta x = \lambda_1 \Delta y$ and $\Delta x = \lambda_2 \Delta y$ are such that $f(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0)$ (to second order)? As above, we need to solve $\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} = 0$ at $(\frac{4}{3}, \frac{4}{3})$: obtaining $2(\lambda^2 + 2\lambda - 1) = 0$ so $\lambda_{1,2} = -1 \pm \sqrt{2}$.

5.1 Characterizing stationary points with the Hessian:

We can write Eq. (5.1) in terms of the Hessian:

$$\Delta x^2 f_{xx} + 2\Delta x\Delta y f_{xy} + \Delta y^2 f_{yy} = \Delta\underline{S}^T H \Delta\underline{S} \tag{5.15}$$

where $\Delta\underline{S} = (\Delta x, \Delta y)$ and here:

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}. \tag{5.16}$$

A local maximum has $\Delta\underline{S}^T H\Delta\underline{S} < 0$ for all $\Delta\underline{S}$. This is the same as saying the matrix H is negative definite (later you'll learn this means its eigenvalues are strictly negative). The condition $f_{xy}^2 - f_{xx}f_{yy} < 0$ can be interpreted as the statement that the determinant of H must be positive. Recall that $\det H = f_{xx}f_{yy} - f_{xy}^2$.

A local minimum has $\Delta\underline{S}^T H\Delta\underline{S} > 0$ for all $\Delta\underline{S}$. This is equivalent to H being positive definite (strictly positive eigenvalues). The conditions for being a minimum are thus $\det H > 0$ and $f_{xx} > 0$.

A saddle point has $\det H < 0$ and H is called indefinite (its eigenvalues have mixed signs).

If $\det H = 0$ higher order terms in the Taylor series are required to characterize the stationary point.

[A notationally involved, and non-examinable, generalisation of the above, for functions of k variables, is that a sufficient condition for a stationary point to be a maximum is to have $(-1)^i D^{(i)} > 0$ for all $i \le k$ where $D^{(i)} = \det H^{(i)}$ and where $H^{(i)}$ is a submatrix of H composed of all entries H_{lm} with indices $l, m < i+1$. $H^{(i)}$ is sometimes called the i^{th} order leading principal minor of H. For a minimum the equivalent condition is that $D^{(i)} > 0$ for all $i \le k$]

6 Leibnitz' Integral Rule

Differentiation of integrals of functions of multiple variables

Consider:

$$F(x) = \int_{t=u(x)}^{t=v(x)} f(x,t)dt \tag{6.1}$$

Leibnitz' Integral Rule is then:

$$\frac{dF(x)}{dx} = f(x, v(x))\frac{dv}{dx} - f(x, u(x))\frac{du}{dx} + \int_{t=u(x)}^{t=v(x)} \frac{\partial f}{\partial x}dt \tag{6.2}$$

Example: If $F(x) = \int_{x^2}^{x^3} \frac{\sin xt}{t} + tdt$ find $\frac{dF}{dx}$. Using Eq. (6.2) we have

$$\frac{dF(x)}{dx} = (\frac{\sin x^4}{x^3} + x^3).3x^2 - (\frac{\sin x^3}{x^2} + x^2).2x + \int_{x^2}^{x^3} \cos xt \, dt \tag{6.3}$$

Where the integral on the right-handside becomes $\left[\frac{\sin xt}{x}\right]_{x^2}^{x^3}$.