EE2 Maths: Stationary Points

Univariate case: When we maximize $f(x)$ (solving 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:

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

Maxima: At a maximum *all* small steps away from (x_0, y_0) lead to $\Delta f < 0$. From Eq. (1) it follows that for a maximum:

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

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

$$f_{xx} < 0 \tag{3}$$

(also $f_{yy} < 0$ by similar arguments). Since Eq. (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. (2) becomes:

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} < 0. \tag{4}$$

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

$$\lambda^2 f_{xx}^2 + 2\lambda f_{xx} f_{xy} + f_{xx} f_{yy} > 0 \tag{5}$$

$$f_{xy}^2 - f_{xx} f_{yy} < (f_{xx}\lambda + f_{xy})^2, \tag{6}$$

for all λ (this last step is a useful trick). The right-hand side of Equation (6) can be set to zero but can't be smaller so:

$$f_{xy}^2 - f_{xx} f_{yy} < 0. \tag{7}$$

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

$$\begin{aligned} f_{xx} &> 0 \qquad (8) \\ f_{xy}^2 - f_{xx}f_{yy} &< 0. \qquad (9) \end{aligned}$$

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

Saddle Points: In the Taylor expansion for Δf what happens if $\Delta x = \lambda \Delta y$ such that Eq. (1) is zero to second order?

$$\lambda^2 f_{xx} + 2\lambda f_{xy} + f_{yy} = 0. \qquad (10)$$

We can use the same trick of multiplying through by f_{xx} and rearranging as we did in Eqs. (5,6) and noting that $(f_{xx}\lambda + f_{xy})^2 > 0$ for all λ we find:

$$f_{xy}^2 - f_{xx}f_{yy} > 0 \qquad (11)$$

which is the condition for a stationary point to be a saddle point. Note that f_{xx}, f_{yy} are unconstrained. Eq. (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.

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

$$\begin{aligned} Max: \quad f_{xx} < 0 \quad \text{and} \quad f_{xy}^2 - f_{xx}f_{yy} &< 0 \qquad (12) \\ Min: \quad f_{xx} > 0 \quad \text{and} \quad f_{xy}^2 - f_{xx}f_{yy} &< 0 \qquad (13) \\ Saddle: \quad f_{xy}^2 - f_{xx}f_{yy} &> 0 \qquad (14) \end{aligned}$$

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. (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 $\lambda_{1,2} = -1 \pm \sqrt{2}$.

Characterizing stationary points with the Hessian:

We can write Eq. (1) in terms of the Hessian:

$$\Delta x^2 f_{xx} + 2\Delta x \Delta y f_{xy} + \Delta y^2 f_{yy} = \Delta \underline{S}^T H \Delta \underline{S} \tag{15}$$

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

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}. \tag{16}$$

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 $i \leq k$ 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 $i \leq k$]

Leibnitz' Integral Rule:
Differentiation of integrals of functions of multiple variables

Consider:

$$F(x) = \int_{t=u(x)}^{t=v(x)} f(x,t)dt \tag{17}$$

Leibnitz' Integral Rule is then:

$$\frac{dF(x)}{dx} = f(x,v(x))\frac{dv}{dx} - f(x,u(x))\frac{du}{dx} + \int_{t=u(x)}^{t=v(x)} \frac{\partial f}{\partial x}dt \tag{18}$$

Example: If $F(x) = \int_{x^2}^{x^3} \frac{\sin xt}{t} + tdt$ find $\frac{dF}{dx}$. Using Eq. (18) we have

$$\frac{dF(x)}{dx} = (\frac{\sin x^4}{x^3} + x^3).3x^2 - (\frac{\sin x^3}{x^2} + x^2).2x + \int_{x^2}^{x^3} \cos xt\, dt \tag{19}$$

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