LECTURE 23 : 2nd Derivative Test

- pgs 252 - 257 of 2020 Lecture Notes.

$f(x,y)$ considering critical point (x_0, y_0) where $\nabla f(x_0, y_0) = \langle 0, 0 \rangle$

$$f(x,y) = f(x_0, y_0) + \underbrace{Q(x - x_0, y - y_0)} + \cdots$$

λ_1, λ_2 : eigenvalues of A

$$Q(x-x_0, y-y_0) = [x-x_0, y-y_0] \underbrace{\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}}_{A} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} = \lambda_1 \bar{x}^2 + \lambda_2 \bar{y}^2$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \det(A) = ad - bc = \lambda_1 \lambda_2 > 0 \text{ for min/max}$$

$$\text{trace}(A) = a + d = \lambda_1 + \lambda_2 \quad + \text{(min)}, \; - \text{(max)}$$

$$\det[Q] = \det(A) = \underbrace{f_{xx} f_{yy} - f_{xy}^2}_{\text{"D"}} \quad // \; tr(A) = f_{xx} + f_{yy}$$

5.3 critical point analysis

Let's focus on the $n = 2$ case since that is the only case we can work out in general[4]. If $f : dom(f) \subseteq \mathbb{R}^2 \to \mathbb{R}$ and (x_o, y_o) is a critical point then $f_x(x_o, y_o) = f_y(x_o, y_o) = 0$ hence the taylor expansion at (x_o, y_o) provides the following representation of f:

$$f(x_o + h, y_o + k) = f(x_o, y_o) + \underbrace{\frac{1}{2}f_{xx}(x_o, y_o)h^2 + f_{xy}(x_o, y_o)hk + \frac{1}{2}f_{yy}(x_o, y_o)k^2}_{Q(h,k)} + T$$

where T is the tail of the series which factors in higher-derivative corrections. We worked out the general behaviour of a quadratic form in a previous section. Let me quote the result here: The graph of $z = Q(h, k) = ah^2 + 2bhk + ck^2$ for some constants $a, b, c \in \mathbb{R}$ can be categorized by real solutions λ_1, λ_2 of the **characteristic equation** $\lambda^2 - (a + c)\lambda + ac - b^2 = 0$. In particular, if $\lambda_1 \leq \lambda_2$ then $\lambda_1 R^2 \leq Q(h, k) \leq \lambda_2 R^2$ for all (h, k) on the circle $h^2 + k^2 = R^2$. We identify that

$$a = \frac{1}{2}f_{xx}(x_o, y_o), \quad b = \frac{1}{2}f_{xy}(x_o, y_o), \quad c = \frac{1}{2}f_{yy}(x_o, y_o).$$

Let's reason through the cases. If both λ_1, λ_2 share the same sign then we can be sure that $|Q(h, k)| >> |T|$ since T depends on third and higher order powers of the coordinates which are much smaller than quadratic powers near the origin. It follows that $0 < \lambda_1 \leq \lambda_2$ imply $f(x_o, y_o)$ is a local minimum of f. Likewise, $\lambda_1 \leq \lambda_2 < 0$ imply $f(x_o, y_o)$ is a local maximum of f. On the other hand, if $\lambda_1 < 0 < \lambda_2$ then the values of Q are sure to increase and decrease near the point of tangency in such a way that T cannot possibly squelch the behaviour and we find $f(x_o, y_o)$ is not a local extremum. The case $\lambda_1 = 0$ is not as useful since the contributions of T are dominant in the direction associated to $\lambda_1 = 0$, we could find a saddle or a minimum or maximum in such a case, so the final two cases in Theorem 5.1.13 are silenced by the tailed beast.

Put all of this together and we have a generalization of the second derivative test for functions of two variables! We need to work out the formulas for λ_1, λ_2 in our current context to make it useful. Solutions of the quadratic equation $\lambda^2 - (a + c)\lambda + ac - b^2 = 0$ are given by

$$\lambda = \frac{a + c \pm \sqrt{(a + c)^2 - 4(ac - b^2)}}{2}$$

We have $a = \frac{1}{2}f_{xx}(x_o, y_o)$, $b = \frac{1}{2}f_{xy}(x_o, y_o)$, $c = \frac{1}{2}f_{yy}(x_o, y_o)$. To reduce clutter, drop the (x_o, y_o) for the next few computations, the two's in a, b, c nicely cancel with the quadratic formula to yield:

$$\lambda_{\pm} = f_{xx} + f_{yy} \pm \sqrt{(f_{xx} + f_{yy})^2 - 4(f_{xx}f_{yy} - f_{xy}^2)}$$

Let $D = f_{xx}f_{yy} - f_{xy}^2$. We have a few cases to consider:

1. If $D < 0$ then clearly

$$|f_{xx} + f_{yy}| = \sqrt{(f_{xx} + f_{yy})^2} < \sqrt{(f_{xx} + f_{yy})^2 - 4(f_{xx}f_{yy} - f_{xy}^2)}.$$

   This inequality indicates that the radical dominates the sign of the solution; given $D < 0$ we have $\lambda_- < 0$ and $\lambda_+ > 0$. Hence, the condition $D < 0$ signifies a saddle shape for $graph(f)$.

[4] I will work special cases in the $n = 3$ case, but the general problem is too hard w/o the help of linear algebra

2. If $D > 0$ then clearly

$$|f_{xx} + f_{yy}| = \sqrt{(f_{xx} + f_{yy})^2} < \sqrt{(f_{xx} + f_{yy})^2 - 4(f_{xx}f_{yy} - f_{xy}^2)}.$$

   This inequality indicates that $f_{xx} + f_{yy}$ dominates the sign of the solution; in particular:

   (a) if $f_{xx} + f_{yy} > 0$ then $\lambda_\pm > 0$ hence f attains a local minimum at (x_o, y_o)

   (b) if $f_{xx} + f_{yy} < 0$ then $\lambda_\pm < 0$ hence f attains a local maximum at (x_o, y_o)

3. If $D = 0$ then either $\lambda_+ = 0$ or $\lambda_- = 0$ hence the quadratic data is inconclusive. The function may attain a maximum, minimum, a saddle or a trough at the critical point.

Notice that in Case (2.) we can simplify the criteria a bit. If $D > 0$ then $f_{xx}f_{yy} - f_{xy}^2 > 0$ thus $0 \leq f_{xy}^2 < f_{xx}f_{yy}$. It follows that either both f_{xx} and f_{yy} are positive or both are negative. Therefore, given $D > 0$, the criteria $f_{xx} + f_{yy} > 0$ can be replaced with criteria $f_{xx} > 0$ or $f_{yy} > 0$. Likewise, given $D > 0$, the criteria $f_{xx} + f_{yy} < 0$ can be replaced with criteria $f_{xx} < 0$ or $f_{yy} < 0$.

Let us collect these thoughts for future use.

Theorem 5.3.1.

Suppose $f : dom(f) \subseteq \mathbb{R}^2 \to \mathbb{R}$ is smooth and (x_o, y_o) is a critical point of f.
If $D = f_{xx}(x_o, y_o)f_{yy}(x_o, y_o) - f_{xy}(x_o, y_o)^2$ then

1. $D < 0$ implies $f(x_o, y_o)$ is not a local extrema,
2. $D > 0$ and $f_{xx}(x_o, y_o) > 0$ (or $f_{yy}(x_o, y_o) > 0$) implies $f(x_o, y_o)$ is a local minimum,
3. $D > 0$ and $f_{xx}(x_o, y_o) < 0$ (or $f_{yy}(x_o, y_o) < 0$) implies $f(x_o, y_o)$ is a local maximum.

Example 5.3.2. *Suppose $f(x,y) = x^2 + 2xy + 2y^2$ then $\nabla f = \langle 2x + 2y, 2x + 4y \rangle$. The origin $(0,0)$ is a critical point since $\nabla f(0,0) = \langle 0, 0 \rangle$. Let's use the theorem to test what type of critical point we've found. We should calculate all the second derivatives,*

$$f_{xx} = 2, \quad f_{xy} = 2, \quad f_{yy} = 4.$$

Calculate $D = f_{xx}f_{yy} - f_{xy}^2 = 8 - 4 = 4)0$ and note $f_{xx} = 2 > 0$ hence $f(0,0)$ is a local minimum. The graph $z = f(x,y)$ opens upward at the origin.

Example 5.3.3. *Suppose $f(x,y) = -x^2 + 2xy - 2y^2$ then $\nabla f = \langle -2x + 2y, 2x - 4y \rangle$. The origin $(0,0)$ is a critical point since $\nabla f(0,0) = \langle 0, 0 \rangle$. Let's use the theorem to test what type of critical point we've found. We should calculate all the second derivatives,*

$$f_{xx} = -2, \quad f_{xy} = 2, \quad f_{yy} = -4.$$

Calculate $D = f_{xx}f_{yy} - f_{xy}^2 = 8 - 4 = 4 > 0$ and note $f_{xx} = -2 < 0$ hence $f(0,0)$ is a local maximum. The graph $z = f(x,y)$ opens downward at the origin.

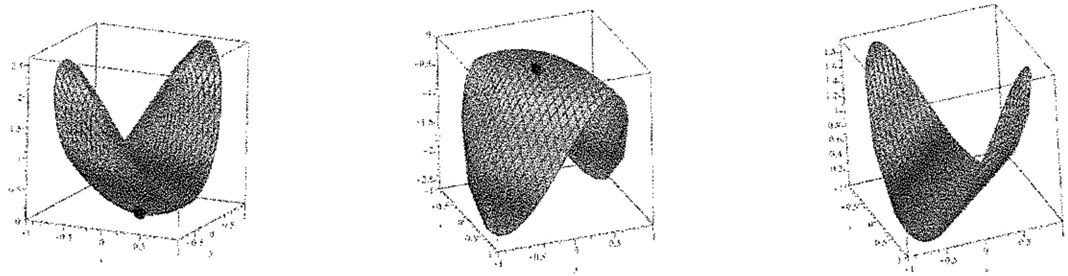
Example 5.3.4. *Suppose $f(x,y) = x^2 + 2xy + y^2$ then $\nabla f = \langle 2x + 2y, 2x + 2y \rangle$. The origin $(0,0)$ is a critical point since $\nabla f(0,0) = \langle 0, 0 \rangle$. Let's use the theorem to test what type of critical point we've found. We should calculate all the second derivatives,*

$$f_{xx} = 2, \quad f_{xy} = 2, \quad f_{yy} = 2.$$

$f(x,y) = x^2 + 2xy + y^2 = (x+y)^2$

Calculate $D = f_{xx}f_{yy} - f_{xy}^2 = 4 - 4 = 0$. *The multivariate second derivative test fails. We can easily see why in this case. Note that the formula for* $f(x,y)$ *factors* $f(x,y) = (x+y)^2$. *The graph* $z = (x+y)^2$ *is zero all along the line* $y = -x, z = 0$ *and it opens like a parabola in planes normal to this line. In other words, this is just* $z = x^2$ *rotated* 45 *degrees around the z-axis. It's a parabolic trough. Notice there are infinitely many critical points in this example.*

Let us constrast the graphs of the past three examples: I plot the graphs over the unit-disk $x^2+y^2 \leq 1$ for Examples 5.3.2, 5.3.3 and 5.3.4 from left to right respective:



Example 5.3.5. *Let* $f(x,y) = x^3 - 12xy + 8y^3$. *Find and classify any local extrema of* f.

Solution: *begin by locating all critical points:*

$$\nabla f(x,y) = \langle 3x^2 - 12y,\ -12x + 24y^2 \rangle = \langle 0, 0 \rangle$$

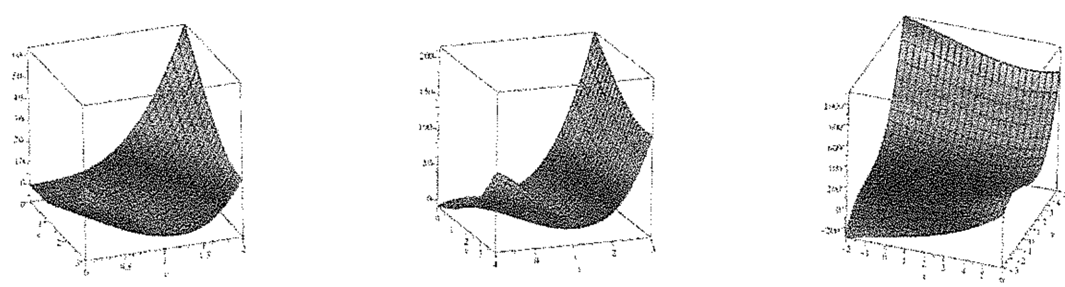
thus, $3x^2 - 12y = 0$ *and* $-12x + 24y^2 = 0$. *Hence,* $4y = x^2$ *and* $x = 2y^2$ *from which we obtain* $4y = (2y^2)^2 = 4y^4$. *Therefore,* $0 = y^4 - y = y(y^3 - 1)$ *hence* $y = 0$ *or* $y = 1$ *and so* $x = 2(0)^2 = 0$ *and* $x = 2(1)^2 = 2$ *respectively. We find critical points* $(0,0)$ *and* $(2,1)$. *The Hessian at* (x,y) *is calculated:*

$$f_{xx} = 6x, \quad f_{yy} = 48y, \quad f_{xy} = -12 \quad \Rightarrow \quad D = 288xy - 144 = 144(2xy - 1).$$

Consider the

critical point	D	f_{xx}	*conclusion*
$(0,0)$	-144	*no need*	*saddle at* $(0,0)$
$(2,1)$	432	12	$f(2,1) = -8$ *is local minimum*

Therefore, we conclude, by the second derivative test, there is only one local extrema. The local minimum value of -8 *is attained at* $(2,1)$. *Graphically, you could easily miss this valley. Consider: the plots below are centered about* $(2,1)$ *and zoom out as you read from left to right:*



It only gets worse as we zoom out further. Thankfully, we need not rely on graphs. I use them to check the answer, not to find it. I think the reader can appreciate why.

Example 5.3.6. *Suppose $f(x,y) = (2x - x^2)(2y - y^2)$. Find and classify any local extrema of f.*

Solution: *we must find the critical points where $\nabla f = 0$. Consider,*

$$\langle f_x, f_y\rangle = \big\langle (2-2x)(2y-y^2), (2x-x^2)(2-2y)\big\rangle$$

Factoring $2, y, x$ reveal:

$$f_x : \;\; 2(1-x)y(2-y) = 0 \qquad \& \qquad f_y : \;\; 2x(2-x)(1-y) = 0$$

We must simultaneously solve the equations above. To solve $f_x = 0$ we have three cases:

(i.) *If $x = 1$ then we require $y = 1$ to solve $f_y = 0$.*

(ii.) *If $y = 0$ then we either need $x = 0$ or $x = 2$ to solve $f_y = 0$.*

(iii.) *If $y = 2$ then we either need $x = 0$ or $x = 2$ to solve $f_y = 0$.*

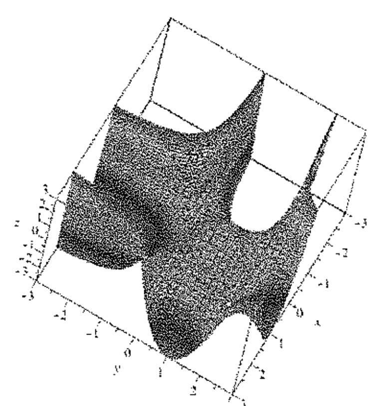
In summary, we find critical points $(1,1), (0,0), (2,0), (0,2), (2,2)$. The Hessian is derived below:

$$\begin{aligned} D(x,y) &= f_{xx}f_{yy} - f_{xy}^2 \\ &= [(-2)(2y-y^2)][(2x-x^2)(-2)] - [(2-2x)(2-2y)]^2 \\ &= 4xy(2-y)(2-x) - 16(1-x)^2(1-y)^2 \end{aligned}$$

Therefore, we find:

critical point	D	f_{xx}	*conclusion*
$(1,1)$	4	-2	$f(1,1) = 1$ *is local maximum*
$(0,0)$	-16	*no need*	*saddle at* $(0,0)$
$(2,0)$	-16	*no need*	*saddle at* $(2,0)$
$(0,2)$	-16	*no need*	*saddle at* $(0,2)$
$(2,2)$	-16	*no need*	*saddle at* $(2,2)$

The plot below uses the "zhue" option to indicate z-values by color. You can clearly see which point is $(1,1)$ and the saddle points are situated symmetrically about the point as our analysis predicted:



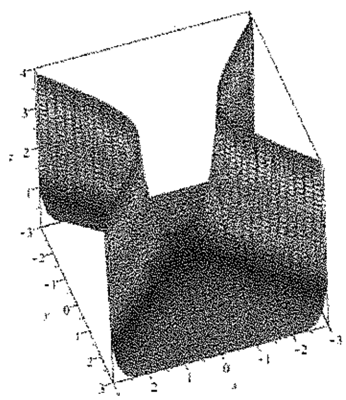
Example 5.3.7. *Suppose $f(x,y) = e^{-x^2+y^2}$ calculate $\nabla f = \langle -2xe^{-x^2+y^2}, 2ye^{-x^2+y^2}\rangle$ and note the origin is the only critical point since exponential functions are strictly positive. Once more we use the multivariate second derivative test at the origin. We need to calculate second derivatives,*

$$f_{xx} = (-2-4x^2)e^{-x^2+y^2}, \quad f_{xy} = -4xye^{-x^2+y^2}, \quad f_{xx} = (2+4y^2)e^{-x^2+y^2}$$

Hence, $f_{xx}(0,0) = -2$, $f_{xy} = 0$ and $f_{yy} = 2$. Note then that

$$D = f_{xx}f_{yy} - f_{xy}^2 = -4 < 0$$

Therefore, $f(0,0)$ is not a local extremum. The graph of $z = f(x,y)$ is saddle shaped over $(0,0)$.



$e^{\theta} = 1 + \theta + \frac{1}{2}\theta^2 + \cdots$

$\theta = -x^2+y^2$

Notice that in the last example it is easy to see why we find the result we did since

$$f(x,y) = e^{-x^2+y^2} = 1 + y^2 - x^2 + \frac{1}{2}(y^2-x^2)^2 + \cdots$$

The fourth order and higher terms are very small compared to the quadratic terms near the origin hence to a good approximation the graph $z = f(x,y)$ looks like $z = 1 + y^2 - x^2$. This is the type of function we can analyze without the help of linear algebra. Let me illustrate by example.

Example 5.3.8. *Suppose $f(x,y,z) = \sin(x^2+y^2+z^2)$ then you can calculate that $\nabla f(0,0,0) = \langle 0,0,0\rangle$ hence the origin is a critical point. Applying the power series expansion for sine,*

$$f(x,y,z) = x^2+y^2+z^2 - \frac{1}{6}(x^2+y^2+z^2)^3 + \cdots$$

clearly $f(0,0,0)$ is a local minimum for f since the values clearly increase. This is clear because the quadratic terms dominate near $(0,0,0)$. On the other hand, if $g(x,y,z) = \sin(x^2+y^2-z^2)$ then

$$g(x,y,z) = x^2+y^2-z^2 - \frac{1}{6}(x^2+y^2-z^2)^3 + \cdots$$

and it is clear that the values of g both increase and decrease near $(0,0,0)$. For example, $g(x,0,0) = x^2 + \cdots$ whereas $g(0,0,z) = -z^2 + \cdots$. It follows that $g(0,0,0)$ is neither a maximum nor a minimum.

The logic used in the example above is not so easy if there are cross terms. For example, $f(x,y,z) = x^2+y^2+z^2+2xy+2xz$ has critical point $(0,0,0)$ but I wouldn't ask you to ascertain the behaviour of f at $(0,0,0)$ because we need linear algebra to understand clearly how f behaves.

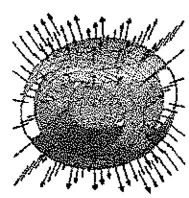
5.3.1 a view towards higher dimensional critical points*

If the function $f : dom(f) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is *analytic* at $\vec{p}$ then that means it is well-approximated by its multivariate Taylor series near $\vec{p}$. For such a function f the statement $\vec{p}$ is a critical point is to say $\nabla f(\vec{p}) = 0$. It follows the taylor series at $\vec{p}$ has the form

$$\boxed{f(\vec{p}+\vec{h}) = f(\vec{p}) + Q(\vec{h}) + T}$$

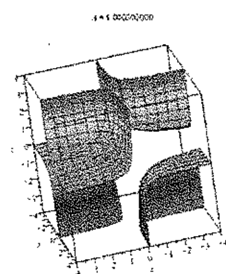
where $|T|$ is usually smaller than $|Q(\vec{h})|$. We call T the **tail** of the expansion. To judge if Q or T dominates the behaviour of f near $\vec{p}$ we must calculate the spectrum of Q. If the spectrum consists of all positive eigenvalues then $f(\vec{p})$ is a local minimum. If the spectrum consists of all negative eigenvalues then $f(\vec{p})$ is a local maximum. If the spectrum consists of both positive and negative eigenvalues then $f(\vec{p})$ is a not a local extrema. If zero is an eigenvalue of Q then further analysis beyond quadratic data may be needed to ascertain the nature of the critical point.

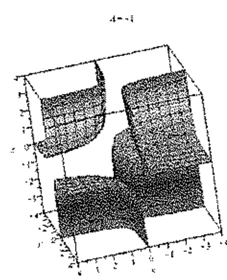
Incidentally, there is a way to visualize maxima for functions of three variables in terms of level surfaces. It's the analogue of using two-dimensional contour plots for finding max/min of a three-dimensional graph. For example, the function $f(x, y, z) = x^2 + 2y^2 + 3z^2$ has level surfaces which are ellipsoids centered at the origin.



You can *see* how the ellipsoids enfold the origin. The larger ellipsoids correspond to higher levels and there does not exist a negative level surface. Intuitively it is clear that $f(0, 0, 0) = 0$ is a local minimum of the function f near the origin. I don't teach this as a method because few of us are capable of mastering such visualization with any reliability. On the other hand, contour plots are extremely useful because our minds are much more adept at handling two-dimensional data.

Consider $f(x, y, z) = xyz$. The origin $(0, 0, 0)$ is a critical point. Plotted below are the level surfaces $xyz = 1$ and $xyz = -1$. In this case $f(0, 0, 0)$ is not a local extreme.





Intuitively, if we have a critical point where f has a trivial quadratic term and a nontrivial cubic term then I expect it is not a local extreme. On the other hand, if the first nontrivial term beyond the constant term is fourth order then max/min or saddle-type points ought to exist. For example,

$$f(x, y) = x^2y^2, \quad f(x, y) = -x^2y^2, \quad f(x, y) = xy^3.$$