13.10. Interior critical points of functions of two variables

Previously, we have concluded that if \( f \) has derivatives, all \textbf{interior} local min or local max should be critical points. Do we have a way to distinguish which are local max and which are local min?

Actually, we do have some tools to achieve this goal to some extent. Now, we study a method to classify the interior critical points of a function of two variables: \textbf{Second derivative test}. \textit{Note that this method may be unable to classify some points.}

If \((a, b)\) is a critical point, \( \nabla f(a, b) = 0 \).

The Taylor expansion to second order reads:

\[
f(x, y) = f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b) + \frac{1}{2} f_{xx}(a, b)(x-a)^2 + f_{xy}(a, b)(x-a)(y-b) + \frac{1}{2} f_{yy}(a, b)(y-b)^2 + \text{even smaller}
\]

Whether the critical point is a local max, local min, or a saddle, might be determined by the \textbf{quadratic form}

\[
q(h, k) = \frac{1}{2} f_{xx}(a, b)h^2 + f_{xy}(a, b)hk + \frac{1}{2} f_{yy}(a, b)k^2
\]

- If \( f_{xx} \neq 0 \), we can complete the square:

\[
q = \frac{1}{2 f_{xx}} \left[ (f_{xx}h + f_{xy}k)^2 + (f_{xx}f_{yy} - f_{xy}^2)k^2 \right]
\]

Clearly, \( f_{xx}f_{yy} - f_{xy}^2 > 0 \), then \( q = \frac{1}{2 f_{xx}} (u^2 + v^2) \) has a definite sign for \((h, k) \neq 0\). Hence, the form is definite.

- If \( f_{xx} > 0 \), then \( q = \frac{1}{2 f_{xx}} (u^2 + v^2) \) is positive but 0 at \((h, k) = (0, 0)\). The point \((a, b)\) is a local min.

- If \( f_{xx}f_{yy} - f_{xy}^2 > 0 \) but \( f_{xx} < 0 \), \( q = \frac{1}{2 f_{xx}} (u^2 + v^2) \) is negative but 0 at \((h, k) = (0, 0)\). The function has a local max.
If $f_{xx}f_{yy} - f_{xy}^2 < 0$, $q = \frac{1}{2f_{xx}}(u^2 - v^2)$, then the point is a saddle point.

If $f_{xx}f_{yy} - f_{xy}^2 = 0$, then $q = \frac{1}{2f_{xx}}u^2$ is either nonpositive or nonnegative. Along the line determined by $u = 0$, $q$ is a constant zero. This means $(u,v) = (0,\text{anything})$ is a local extreme of the quadratic form. However, the small error may perturb the function a little bit up or a little bit lower and the behavior of $f$ is unclear. The test fails.

- If $f_{xx} = 0$, $f_{xx}f_{yy} - f_{xy}^2 = -f_{xy}^2$. If $f_{xy} \neq 0$, the point is a saddle. If $f_{xy} = 0$, the whole quantity is zero and $q = \frac{1}{2}f_{yy}k^2$. We can say nothing since the small error may perturb.

**In summary**, we can check the sign of $\Delta = f_{xx}f_{yy} - f_{xy}^2$ first.

1. $\Delta > 0$, $f_{xx} > 0$ implies local min and $f_{xx} < 0$ implies local max. ($f_{xx}$ can’t be zero in this case)

2. $\Delta < 0$, saddle.

3. $\Delta = 0$, the test is unable to classify the points. We must use other techniques.

**Exercise**: In the Monkey saddle point case, there are three up branches and three down branches. Which case does the Monkey saddle point belongs to? ($\Delta = 0$ case.)

**Example**: Find and classify the critical points of $f(x,y) = x^2 - 2xy + y^3 - y$.

**Solution.** Set $\nabla f = 0$. We have

$$f_x = 2x - 2y = 0,$$
$$f_y = -2x + 3y^2 - 1 = 0.$$

From the first equation $y = x$, then

$$-2y + 3y^2 - 1 = 0.$$

By using either quadratic formula or factoring it to $(3y + 1)(y - 1) = 0$, we have $y = 1, y = -1/3$. Then, we have $(1,1)$ and $(-1/3,-1/3)$. We can verify that both of them are critical points.

$$f_{xx} = 2, f_{xy} = -2, f_{yy} = 6y.$$
At \((1,1)\),
\[
    f_{xx} = 2, \quad f_{xy} = -2, \quad f_{yy} = 6. \quad \Delta = f_{xx}f_{yy} - f_{xy}^2 = 12 - (-2)^2 > 0.
\]
Since \(f_{xx} > 0\), the point \((1,1)\) is a local min.

At \((-1/3,-1/3)\),
\[
    f_{xx} = 2, \quad f_{xy} = -2, \quad f_{yy} = -2. \quad \Delta = f_{xx}f_{yy} - f_{xy}^2 = -4 - (-2)^2 < 0.
\]
\((-1/3,-1/3)\) is a saddle point. \(\square\)

**Example** Find and classify the critical points of \(z = \sin(x)\sin(y)\) over the region \((-\pi, \pi) \times (-\pi/2, \pi)\).

### 13.9. Constrained optimization and Lagrange multipliers

Review: \(\nabla f\) is the fastest increasing direction and \(\nabla f\) is perpendicular with the level set.

If \(u\) is not along \(\nabla f\) but it has an acute angle from \(\nabla f\), moving along \(u\), we see that \(f\) increases.

We study a method for finding the maximum and minimum over a level set of another function on which there are no interior points. On such a structure, the gradient at the extremum does not have to zero.

**Example:** Find the point on the surface \(xyz = 1, x > 0, y > 0, z > 0\) such that \(f(x,y,z) = 2x^2 + y^2 + z^2\) achieves the smallest value.

**Solution.** In the 3D space, the surface \(xyz = 1\) contains no interior point. Hence, at minimum point, \(\nabla f\) may not be 0. One idea is to eliminate one variable, for example \(z\) and have \(h(x,y) = 2x^2 + y^2 + \frac{1}{x^2y^2}\).

Now, \(h\) as a function of two variables, the region we consider has interior points now (of course in 2D plane instead of 3D space).

As \((x,y)\) is close to \(x, y\) axis or as \((x,y) \to \infty\), \(h\) blows up. Then, there must be a minimum point in the interior. That has to be a critical point of \(h\), or \(\nabla h = 0\).

\[
    4x - \frac{2}{x^3y^2} = 0
\]
\[
    2y - \frac{1}{x^2y^3} = 0
\]
\[
    x^6 = \frac{1}{4} \quad \text{and} \quad y^4 = \frac{1}{x^2} = \sqrt[3]{4}.
\]
\(\square\)
The method of Lagrange multipliers

Usually, $z$ is hard to solve. Then, we use another method called the method of Lagrange multipliers.

With one constraint

The method is given by the following theorem:

**Theorem 1.** Suppose $f$ and $g$ are two continuously differentiable functions. On the level set $g = C$ (or in other words, with constraint $g = C$), if the maximum (or minimum) of $f$ happens at $P$, and $\nabla g \neq 0$, then $\nabla f$ and $\nabla g$ are parallel or

$$\nabla f = \lambda \nabla g,$$

for some $\lambda$. $\lambda$ is called the Lagrange multiplier.

(Draw a picture. If $\nabla f$ is not parallel with $\nabla g$, then along the level set of $g$, $f$ will increase. Also, consider drawing the level sets of the two functions and show the idea.)

**Proof.** Let $r(t)$ be any curve in the level set such that $r(t_0) = P$. Then, the function

$$m(t) = f(r(t)),$$

achieves a local maximum or minimum at $t = t_0$. Hence, $m'(t_0) = \nabla f(P) \cdot r'(t_0) = 0$ by the chain rule. Hence, $\nabla f(P)$ is perpendicular with the curve. Since the curve is arbitrary in the level set, $\nabla f(P)$ must be perpendicular with the whole level set.

On the other hand, $\nabla g(P)$ is perpendicular with the level set. Hence, $\nabla f(P)$ and $\nabla g(P)$ must be parallel. $\square$

To consider possible candidates. We should also consider the cases when $\nabla g = 0$ and the points where the functions are not differentiable.

**Example:** Find the smallest value of $f(x, y, z) = 2x^2 + y^2 + z^2$ on the surface $g = xyz = 1, x > 0, y > 0, z > 0$ using Lagrange multiplier.

**Solution.** Using the Lagrange Multiplier,

$$\nabla f = \lambda \nabla g$$

$$g = 1.$$
Or

\[ 4x = \lambda yz, \]
\[ 2y = \lambda xz, \]
\[ 2z = \lambda xy, \]
\[ xyz = 1. \]

Since \( yz = 1/x \), the first equation tells us that \( 4x = \lambda/x \) or \( \lambda = 4x^2 \).

Hence, \( 4x^2 = 2y^2 = 2z^2 \).

With the fact \( xyz = 1 \), we have \( x^2y^2z^2 = 1 \) or \( x^2 + 2x^2 + 2x^2 = 1 \) or \( x^6 = 1/4 \) which agrees with what we had. Similarly, we can solve \( y \) and \( z \).

Further, in principle, we should also consider \( \nabla g = 0, g = 1 \). However, since \( xyz = 1, \nabla g \) can’t be zero. Then, we have covered all possibilities. □

Example: Find the points of the rectangular hyperbola \( xy = 1 \) in the first quadrant that are closest to the origin \((0,0)\).

Solution. Let \((x, y)\) be on the rectangular hyperbola. The distance to the origin is \( d = \sqrt{x^2 + y^2} \). Then, we would like to minimize

\[ f(x, y) = d^2 = x^2 + y^2, \]

with constraint \( g(x, y) = xy = 1 \).

Further, since \( xy = 1, \nabla g \) is never 0 on the constraint. Hence, we only need to consider \( \nabla f = \lambda \nabla g \). By the Lagrange multiplier method

\[ f_x = \lambda g_x, f_y = \lambda g_y, g = xy = 1, \]

or

\[ 2x = \lambda y \]
\[ 2y = \lambda x \]
\[ xy = 1 \]

hence, \( 2x^2 = 2y^2 \). For the point in the first quadrant, we have \( x = y \). Hence, \( x^2 = 1 \) or \( x = 1 \). \((1, 1)\) is the point and \( d = \sqrt{1^2 + 1^2} = \sqrt{2} \).

Exercise: can we find the points that are farthest to the origin on the surface using Lagrange multipliers? Why?
An example of the global extrema

Find the global extrema of $f = (x - 1)^2 + y^2 + z^2$ in the region $D = \{(x, y, z) : x^2 + 4y^2 + 9z^2 \leq 36\}$

With two constraints (Omitted)

If we have two constraints, $g = C_1$ and $h = C_2$, then at the max or min $P$, $\nabla f(P)$, $\nabla g(P)$ and $\nabla h(P)$ are co-planar (in the linear algebra language: they are linearly dependent). If $\nabla g$ and $\nabla h$ are not parallel, then

$$\nabla f(P) = \lambda \nabla g(P) + \mu \nabla h(P),$$

for two constants $\lambda$ and $\mu$.

Of course, the critical points may also happen at the points where $\nabla g = 0$ or $\nabla h = 0$ or some of them are not differentiable.

Draw a picture to show the idea. The two surfaces $g = C_1$ and $h = C_2$ intersect at one curve. The curve is perpendicular with both $\nabla g$ and $\nabla h$. If $\nabla f$ is not in the plane determined by $\nabla g$ and $\nabla h$, then $f$ will increase along the curve.

Example: Suppose we want to maximize $f$ on the constraint $g = 1$ and $h = 2$. If at a point $P$ on the intersection of the two surfaces, $\nabla g = \langle 1, 1, 2 \rangle$ and $\nabla h = \langle -1, -2, 3 \rangle$ and $\nabla f = \langle 2, 3, -2 \rangle$. Is it possible for this point to be an extremum?

Example: Find the highest and lowest points on the intersection of $x + y + z = 12$ and $z = x^2 + y^2$.

Solution. Clearly the height is given by the $z$-coordinate.

$$f(x, y, z) = z, \quad g(x, y, z) = x + y + z = 12, \quad h(x, y, z) = x^2 + y^2 - z = 0.$$ 

by the method of Lagrange multiplier,

$$\nabla f = \lambda \nabla g + \mu \nabla h$$

$$\begin{align*}
x + y + z &= 12 \\
x^2 + y^2 - z &= 0
\end{align*}$$
or

\begin{align*}
0 &= \lambda + \mu 2x \\
0 &= \lambda + \mu 2y \\
1 &= \lambda - \mu \\
x + y + z &= 12 \\
x^2 + y^2 - z &= 0
\end{align*}

Hence, \(2\mu(x - y) = 0\). If \(\mu = 0\), the from the first equation, \(\lambda = 0\) which can’t make the third equation to be true. Hence, \(\mu \neq 0\) and \(x = y\). Then, \(2x + z = 12\) and \(2x^2 - z = 0\) which gives \(2x + 2x^2 = 12\) or \(x^2 + x - 6 = 0\), \(x = -3, 2\). We therefore have two points \((-3, -3, 18)\) and \((2, 2, 8)\). The first point is the highest while the second one is the lowest. \qed