Local Extrema

Now that we've defined the Hessian matrix and studied quadratic forms, we're in position to find local extrema of multivariable function. This process will closely resemble the process that we following in single variable calculus, so we'll begin by reviewing that case.

Recall that we say $f(x)$ has a local maximum at $x = a$ if $f(x) \leq f(a)$ for all x near a. Similarly, $f(x)$ has a local minimum at $x = a$ if $f(a) \leq f(x)$ for all x near a. We can state these definitions more precisely, as below.

Definition 1. *We say that $f(x)$ has a* local maximum *at $x = a$ if there exists $r > 0$ such that $|x - a| < r$ implies $f(x) \leq f(a)$.*

We say that $f(x)$ has a local minimum *at $x = a$ if there exists $r > 0$ such that $|x - a| < r$ implies $f(x) \geq f(a)$.*

In single variable calculus, we found local extrema by finding critical points, and then classifying them using the first or second derivative test.

Definition 2. *The point $x = a$ is a* critical point *of $f(x)$ if $f'(a) = 0$, or if $f'(a)$ does not exist.*

Proposition 1. *If $f(x)$ has a local minimum or maximum at $x = a$, then a is a critical point of $f(x)$.*

If a is a critical point of $f(x)$ and $f''(a) < 0$, then $f(x)$ has a local maximum at $x = a$.

If a is a critical point of $f(x)$ and $f''(a) > 0$, then $f(x)$ has a local minimum at $x = a$.

If a is a critical point of $f(x)$ and $f''(a) = 0$ or $f''(a)$ does not exist, then $f(x)$ could have a local maximum at a, a local minimum at a, or neither.

Example 1. *As an example, we'll find the local extrema of the function $f(x) = x^3 + 3x^2 - 9x + 1$. Differentiating, we have*

$$f'(x) = \boxed{3x^2 + 6x - 9}.$$

Solving $f'(x) = 0$ for x, we see that there are two critical points, $x = -3$ and $x = 1$.

We'll classify these critical points using the second derivative test. We find the second derivative of $f(x)$.

$$f''(x) = \boxed{6x + 6}$$

Learning outcomes: Compute critical points to find local extrema, and use the Hessian matrix to classify them.

Author(s): Melissa Lynn

Plugging in $x = -3$, we have

$$f''(-3) = \boxed{-12}.$$

This tells us that at $x = -3$, $f(x)$ has a

Multiple Choice:

(a) *local maximum.* ✓

(b) *local minimum.*

(c) *neither.*

Plugging in $x = 1$, we have

$$f''(1) = \boxed{12}.$$

This tells us that at $x = 1$, $f(x)$ has a

Multiple Choice:

(a) *local maximum.*

(b) *local minimum.* ✓

(c) *neither.*

Local extrema for multivariable functions

We begin by defining local minima and local maxima for multivariable functions. These follow the same idea as in the single variable case. For example, f has a local minimum at $\vec{x} = \vec{a}$ if $f(\vec{a}) \leq f(\vec{x})$ for $\vec{x}$ "near" a. Now, we need to decide what "near" means. This means that we can find some distance $r > 0$ such that for $\vec{x}$ within distance r from $\vec{a}$, we must have $f(\vec{a}) \leq f(\vec{x})$. We restate this more formally.

Definition 3. *We say that $f : \mathbb{R}^n \to \mathbb{R}$ has a* local minimum *at $\vec{x} = \vec{a}$ if there exists $r > 0$ such that for all $\vec{x}$ with $\|\vec{x} - \vec{a}\| < r$, we have $f(\vec{a}) \leq f(\vec{x})$.*

We say that $f : \mathbb{R}^n \to \mathbb{R}$ has a local maximum *at $\vec{x} = \vec{a}$ if there exists $r > 0$ such that for all $\vec{x}$ with $\|\vec{x} - \vec{a}\| < r$, we have $f(\vec{a}) \geq f(\vec{x})$.*

In some cases, we can determine local extrema using our knowledge about functions, and without using Calculus.

Example 2. *Consider the function $f(x, y) = (x - 1)^2 + (y + 2)^2$. Since the terms $(x - 1)^2$ and $(y + 2)^2$ are always nonnegative, we can see that $f(x, y)$ will always be nonnegative. Since $f(1, -2) = 0$, and $0 \leq f(x, y)$ for all (x, y), this*

means that $f(x, y)$ has a local minimum at $(1, -2)$. (This local minimum is also an absolute maximum.)

*Consider the function $g(x, y, z) = 1 - \sqrt{x^2 + y^2 + z^2}$. Since $\sqrt{x^2 + y^2 + z^2} \geq 0$, we'll always have $g(x, y, z) \leq 1$. Then, since $g(0, 0, 0) = 1$, g has a local maximum at $(0, 0, 0)$.*

Critical points

Our definition of critical points of multivariable functions is also very similar to the definition from single variable calculus. Here, the gradient vector replaces the derivative.

Definition 4. *We say that $f : \mathbb{R}^n \to \mathbb{R}$ has a* critical point *at $\vec{x} = \vec{a}$ if $\nabla f(\vec{a}) = \vec{0}$, or $\nabla f(\vec{a})$ does not exist.*

For functions $f : \mathbb{R}^2 \to \mathbb{R}$, we can visualize critical points as points where the tangent plane is either horizontal, or undefined. Each of the graphs below has a critical point at $(0, 0)$.
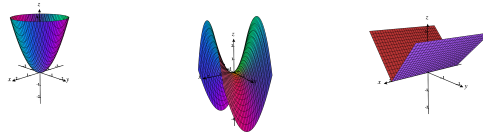
As in single variable calculus, if a function has local extrema, they will occur at critical points.

Proposition 2. *If $f : \mathbb{R}^n \to \mathbb{R}$ has a local maximum or local minimum at $\vec{x} = \vec{a}$, then $\vec{a}$ is a critical point of f.*

Example 3. *We'll find the critical points of the function $f(x, y) = x^3 + x^2 y - y^2 - 4y$. The gradient of f is*

$$\nabla f(x, y) = (3x^2 + 2xy, x^2 - 2y - 4).$$

This is defined at all points in $\mathbb{R}^2$, so the critical points will satisfy $\nabla f(x, y) = (0, 0)$. In order to find the critical points, we solve the system of equations

$$\begin{cases} 3x^2 + 2xy = 0 \\ x^2 - 2y - 4 = 0 \end{cases}.$$

Factoring the first equation, we have $x(3x + 2y) = 0$, giving us the cases $x = 0$ or $3x + 2y = 0$.

If $x = 0$, the second equation gives us $y = -2$. So $(0, -2)$ is a critical point.

If $3x + 2y = 0$, plugging $y = \dfrac{-3}{2}x$ into the second equation gives us $x^2 + 3x - 4 = 0$, so $x = 1$ or $x = -4$. This gives us the critical points $(1, -3/2)$ and $(-4, 6)$.

Thus, the critical points of f are $(0, -2)$, $(1, -3/2)$, and $(-4, 6)$.

Classifying critical points

Analogous to the second derivative test from single variable calculus, we can use the Hessian matrix to classify critical points in some cases.

Proposition 3. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has continuous second order partial derivatives (so f has $\mathcal{C}^2$) at and near a critical point $\vec{a} \in U$.*

- *If $Hf(\vec{a})$ is positive definite, then f has a* local minimum *at $\vec{x} = \vec{a}$.*

- *If $Hf(\vec{a})$ is negative definite, then f has a* local maximum *at $\vec{x} = \vec{a}$.*

- *If $Hf(\vec{a})$ is indefinite, then f has a* saddle point *at $\vec{x} = \vec{a}$.*

Example 4. *Consider again the function $f(x, y) = x^3 + x^2 y - y^2 - 4y$. Earlier, we found the critical points of this function, which are $(0, -2)$, $(1, -3/2)$, and $(-4, 6)$. In order to classify these critical points, we find the Hessian matrix of f.*

$$Hf(x, y) = \begin{pmatrix} \boxed{6x + 2y} & \boxed{2x} \\ \boxed{2x} & \boxed{-2} \end{pmatrix}$$

Plugging in the point $(0, -2)$, we have

$$Hf(0, -2) = \begin{pmatrix} \boxed{-4} & \boxed{0} \\ \boxed{0} & \boxed{-2} \end{pmatrix}.$$

Using Sylvester's Theorem, we can see that $Hf(0, -2)$ is

Multiple Choice:

(a) *positive definite.*

(b) *negative definite.* ✓

(c) *indefinite.*

(d) *we cannot determine this using Sylvester's Theorem.*

This means that, at $(0, -2)$, $f(x, y)$ has a

Multiple Choice:

(a) *local maximum.* ✓

(b) *local minimum.*

(c) *saddle point.*

(d) *none of the above.*

Next, we'll classify the critical point $(1, -3/2)$. We find the Hessian matrix at this point,

$$Hf(1, -3/2) = \begin{pmatrix} \boxed{3} & \boxed{2} \\ \boxed{2} & \boxed{-2} \end{pmatrix}.$$

Using Sylvester's Theorem, we can see that $Hf(1, -3/2)$ is

Multiple Choice:

(a) *positive definite.*

(b) *negative definite.*

(c) *indefinite.* ✓

(d) *we cannot determine this using Sylvester's Theorem.*

This means that, at $(1, -3/2)$, $f(x, y)$ has a

Multiple Choice:

(a) *local maximum.*

(b) *local minimum.*

(c) *saddle point.* ✓

(d) *none of the above.*

Finally, we'll classify the critical point $(-4, 6)$. We find the Hessian matrix at this point,

$$Hf(-4, 6) = \begin{pmatrix} \boxed{-12} & \boxed{-8} \\ \boxed{-8} & \boxed{-2} \end{pmatrix}.$$

Using Sylvester's Theorem, we can see that $Hf(-4, 6)$ is

Multiple Choice:

(a) *positive definite.*

(b) *negative definite.*

(c) *indefinite.* ✓

(d) *we cannot determine this using Sylvester's Theorem.*

This means that, at $(-4, 6)$, $f(x, y)$ has a

Multiple Choice:

(a) *local maximum.*

(b) *local minimum.*

(c) *saddle point.* ✓

(d) *none of the above.*