17. Unconstrained Optimization

10/27/20

A quick summary of this chapter is that

1. A necessary condition for $\mathbf{x}_0$ to be an interior optimal point of the function $f(\mathbf{x})$ is that $Df(\mathbf{x}_0) = \mathbf{0}$. This equation is called the *first order necessary condition*.[1]
2. Appropriate conditions on the second derivative together with the first order necessary condition are sufficient for an optimum

17.1 Maxima and Minima

We will use the terms *max*, *maximizer*, and *maximum point* more or less interchangeably. The same is true of *min*, *minimizer*, and *minimum point*. They refer to points where a function takes its maximum or minimum value. If it is unlikely to cause confusion, we may identify the maximum or minimum point with the maximum or minimum value of the function, both of which can be called the max or min.

Types of Maxima (and Minima). Let $f\colon U \to \mathbb{R}$ be a real-valued function with domain $U \subset \mathbb{R}^m$.

- A point $\mathbf{x}^* \in U$ is a *max* or *global max* of f over U if $f(\mathbf{x}^*) \ge f(\mathbf{x})$ for all $\mathbf{x} \in U$.
- A point $\mathbf{x}^* \in U$ is a *strict max* or *strict global max* of f over U if $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \ne \mathbf{x}^*$ with $\mathbf{x} \in U$.
- A point $\mathbf{x}^* \in U$ is a *local max* or *relative max* of f if there is a ball $B_\varepsilon(\mathbf{x}^*)$ so that $f(\mathbf{x}^*) \ge f(\mathbf{x})$ for all $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$.
- A point $\mathbf{x}^* \in U$ is a *strict local max* or *strict relative max* of f if there is a ball $B_\varepsilon(\mathbf{x}^*)$ so that $f(\mathbf{x}^*) > f(\mathbf{x})$ for all $\mathbf{x} \ne \mathbf{x}^*$ with $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$.

Similar terminology applies to minimum points.

If the graph of a function has several hills, the tops of hill represent local maxima. The top of the highest hill is a global maximum.

[1] When $Df = \mathbf{0}$ involves more than one equation, we sometimes refer to *first order necessary conditions*.

17.2 First Order Conditions

As with functions from $\mathbb{R}$ to $\mathbb{R}$, our main technique for finding optima will start by rounding up the usual suspects. In the case of optimization, the usual suspects are the *critical points*, points $\mathbf{x}^*$ where $Df(\mathbf{x}^*) = \mathbf{0}$. Critical points are not guaranteed to be optimal, but all of the optimal points will be critical points. For interior optimal points, it is **necessary** that they be critical points, but this is **not sufficient** to show they are in fact optima. See the following example.

▶ Example 17.2.1: Non-Optimal Critical Point. Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$. Then $f'(x) = 3x^2$, so $x = 0$ is the only critical point. It is neither a max nor a min because for any $x > 0$, $f(x) > 0 = f(0)$ and for any $x < 0$, $f(x) < 0$. ◀

Theorem 17.2.2 establishes that all interior optima are critical points.

Theorem 17.2.2. *Let $f\colon U \to \mathbb{R}$ be a $\mathcal{C}^1$ function defined on an subset $U \subset \mathbb{R}^m$. If $\mathbf{x}^*$ is a local max or local min of f in U and an interior point of U, then $Df(\mathbf{x}^*) = \mathbf{0}$. That is,*

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0 \quad \textit{for } i = 1, \dots, m.$$

Proof. Consider the difference quotient. We will consider the case of a maximizer $\mathbf{x}^*$. Then $f(\mathbf{x}^* + h\mathbf{e}_i) \le f(\mathbf{x}^*)$ for all $h \neq 0$ and $i = 1, \dots, m$. It follows that

$$\begin{aligned} \frac{f(\mathbf{x}^* + h\mathbf{e}_i) - f(\mathbf{x}^*)}{h} &\le 0 \qquad \text{when } h > 0 \\ \frac{f(\mathbf{x}^* + h\mathbf{e}_i) - f(\mathbf{x}^*)}{h} &\ge 0 \qquad \text{when } h < 0. \end{aligned}$$

Taking the limit as $h \to 0$, we find

$$0 \le \frac{\partial f}{\partial x_i}(\mathbf{x}^*) \le 0.$$

This holds for all $i = 1, \dots, m$, so

$$\frac{\partial f}{\partial x_i}(\mathbf{x}^*) = 0 \quad \text{for } i = 1, \dots, m.$$

∎

There are weaker versions of this theorem that apply at boundary points. For some intuition about this, consider an interval $[a, b] \subset \mathbb{R}$. A differentiable function with a local max at a will have $f'(a) \le 0$ while a local max at b will have $f'(b) \ge 0$. Strict local maxima will have $f'(a) < 0$ and $f'(b) > 0$.

17.3 Toward the Second Order Conditions

We can use the Theorem 17.2.2 to identify possible optimal points. We need more information to determine whether such points are optima.

We turn to Taylor's formula for help in that effort. Let $\mathbf{x}^*$ be a critical point, then there is a point $\mathbf{y}$ in $\ell(\mathbf{x}, \mathbf{x}^*)$ with

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^*) + Df(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}} \left[D^2 f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*) \\ &= f(\mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}} \left[D^2 f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*) \end{aligned} \tag{17.3.1}$$

because $Df(\mathbf{x}^*) = \mathbf{0}$.

We can rewrite this as

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}} \left[D^2 f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*) \tag{17.3.2}$$

The left-hand side of equation (17.3.2) will always be non-positive if the critical point $\mathbf{x}^*$ is a maximum and non-negative if $\mathbf{x}^*$ is a mimimum. If the optimum is strict, that changes to negative for a strict minimum and positive for a strict maximum.

Unless it is zero, the behavior of the second derivative will determine whether a critical point is a maximum or minimum.

If the second derivative is zero, we need to check higher order terms, if possible. To get a little feel for that, if both first and second derivatives are zero, a non-zero third derivative ensures that an interior critical point is neither max nor min. If the third derivative is zero, we then check the fourth derivative to see if it is positive or negative definite. If it is zero, we continue to the fifth derivative, etc.

An example is $f(x) = x^4$. This has a minimum at $x = 0$. Now $f'(x) = 4x^3$, yielding $x = 0$ as a critical point. However, $f''(0) = 0$, so the second order conditions are not helpful. The third derivative is $f'''(x) = 24x$, which is again 0 at $x = 0$. Finally, the fourth derivative is $f^{(4)}(x) = 2 > 0$, indicating a minimum at $x = 0$.

On $\mathbb{R}^m$, we can consider the tensors of the form

$$F(\mathbf{z}) = D^4 f(\mathbf{x})(\mathbf{z} \otimes \mathbf{z} \otimes \mathbf{z} \otimes \mathbf{z} \otimes \mathbf{z}),$$

which are positive definite if $F(\mathbf{z}) \geq 0$ and $F(\mathbf{z}) = 0$ if and only if $\mathbf{z} = \mathbf{0}$.

17.4 Global Sufficient Conditions: Max and Min

We turn to the global sufficient conditions next because they exhibit the basic argument in its purest form.

Suppose f is defined on U and $\mathbf{x}^*$ is an interior critical point. To use equation (17.3.2), we will need $\ell(\mathbf{x}, \mathbf{x}^*) \subset U$ for $\mathbf{x} \in U$. The weakest condition that ensures this always happens is that the domain U of f is star-shaped with respect to $\mathbf{x}^*$. After all, U being star-shaped with respect to $\mathbf{x}^*$ requires precisely that $\ell(\mathbf{x}, \mathbf{x}^*) \subset U$ for all $\mathbf{x} \in U$.

Of course, we also need $\mathbf{x}^*$ in the interior of U to be sure that optima are critical points, the basis for equation (17.3.2). With that in mind, we can now state a theorem describing sufficient conditions for a maximum or minimum.

Theorem 17.4.1. *Let $f\colon U \to \mathbb{R}$ be a $\mathcal{C}^2$ function defined on an subset $U \subset \mathbb{R}^m$. Suppose $\mathbf{x}^*$ is a critical point of f in the interior of U and that U is star-shaped with respect to $\mathbf{x}^*$.*

1. *If the Hessian $D^2f(\mathbf{x})$ is negative semidefinite for every $\mathbf{x} \in U$ then $\mathbf{x}^*$ maximizes f over U.*
2. *If the Hessian $D^2f(\mathbf{x})$ is positive semidefinite for every $\mathbf{x} \in U$ then $\mathbf{x}^*$ minimizes f over U.*

Proof. We start with equation (17.3.2):

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}} \left[D^2 f(\mathbf{y})\right] (\mathbf{x} - \mathbf{x}^*). \tag{17.3.2}$$

where $\mathbf{y} \in \ell(\mathbf{x}^*, \mathbf{x})$. The line segment $\ell(\mathbf{x}^*, \mathbf{x}) \subset U$ because U is star-shaped with respect to $\mathbf{x}^*$. It follows that $D^2f(\mathbf{y})$ is negative semidefinite, so for all $\mathbf{x} \in U$,

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}^*) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}} \left[D^2 f(\mathbf{y})\right] (\mathbf{x} - \mathbf{x}^*) \\ &\leq 0 \end{aligned}$$

showing that $f(\mathbf{x}) \leq f(\mathbf{x}^*)$ for every $\mathbf{x} \in U$. The point $\mathbf{x}^*$ maximizes f over U.

The proof of case (2) is the same, except that $D^2f(\mathbf{y})$ is positive semidefinite, making $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in U$. ∎

17.5 Global Sufficient Conditions: Strict Max and Min

For strict optima, we need definite Hessians, not semidefinite Hessians. The proofs of cases (1) and (2) are barely modified from cases (1) and (2) in Theorem 17.4.1. Case (3) is only sufficient in a negative sense. It shows that the critical point $\mathbf{x}^*$ cannot be either a max or min. Proving it takes some extra work. We need to use the fact that $\mathbf{x}^*$ is an interior point to shrink down two vectors where the quadratic term of equation (17.3.2) takes opposite signs. This yields two vectors that are actually in U where the quadratic term takes opposite signs. We need them in U so that the corresponding $\mathbf{y}$ where D^2f is evaluated is also in U.

Theorem 17.5.1. *Let $f\colon U \to \mathbb{R}$ be a $\mathcal{C}^2$ function defined on an subset $U \subset \mathbb{R}^m$. Suppose $\mathbf{x}^*$ is a critical point of f in the interior of U and that U is star-shaped with respect to $\mathbf{x}^*$.*

1. *If the Hessian $D^2f(\mathbf{x})$ is negative definite for $\mathbf{x} \in U$ then $\mathbf{x}^*$ strictly maximizes f over U.*
2. *If the Hessian $D^2f(\mathbf{x})$ is positive definite for $\mathbf{x} \in U$ then $\mathbf{x}^*$ strictly minimizes f over U.*
3. *If the Hessian $D^2f(\mathbf{x}^*)$ is indefinite then $\mathbf{x}^*$ is neither a maximizer nor minimizer of f over U.*

You'll notice that part (3) of the theorem doesn't involve a global condition on the Hessian. We only need to know the Hessian at $\mathbf{x}^*$ to show that $\mathbf{x}^*$ is neither a maximum nor a minimum.

We call $\mathbf{x}^*$ a *saddlepoint* of f if the Hessian D^2f is indefinite at $\mathbf{x}^*$. Saddlepoints are neither local maxima nor minima. There are directions where f increases as you move away from $\mathbf{x}^*$ and there are directions where it decreases. If the function maps $\mathbb{R}^2 \to \mathbb{R}$, as is the case $f(\mathbf{x}) = x_1^2 - x_2^2$, the graph is similar to a saddle shape in the vicinity of $\mathbf{x}^*$. In higher dimensions, the shapes can be more complex, but there will still be two-dimensional slices where it looks locally like a saddle.

17.6 Proof of Theorem 17.5.1

Theorem 17.5.1. *Let $f\colon U \to \mathbb{R}$ be a $\mathcal{C}^2$ function defined on an subset $U \subset \mathbb{R}^m$. Suppose $\mathbf{x}^*$ is a critical point of f in the interior of U and that U is star-shaped with respect to $\mathbf{x}^*$.*

1. *If the Hessian $D^2f(\mathbf{x})$ is negative definite for $\mathbf{x} \in U$ then $\mathbf{x}^*$ strictly maximizes f over U.*
2. *If the Hessian $D^2f(\mathbf{x})$ is positive definite for $\mathbf{x} \in U$ then $\mathbf{x}^*$ strictly minimizes f over U.*
3. *If the Hessian $D^2f(\mathbf{x}^*)$ is indefinite then $\mathbf{x}^*$ is neither a maximizer nor minimizer of f over U.*

Proof. We again start with equation (17.3.2):

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \left[D^2f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*). \tag{17.3.2}$$

where $\mathbf{y} \in \ell(\mathbf{x}^*, \mathbf{x})$. The line segment $\ell(\mathbf{x}^*, \mathbf{x}) \subset U$ because U is star-shaped with respect to $\mathbf{x}^*$. It follows that $D^2f(\mathbf{y})$ is negative definite, so for all $\mathbf{x} \in U$ with $\mathbf{x} \neq \mathbf{x}^*$,

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}^*) &= \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \left[D^2f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*) \\ &< 0 \end{aligned}$$

showing that $f(\mathbf{x}) < f(\mathbf{x}^*)$ for every $\mathbf{x} \in U$ other than $\mathbf{x} = \mathbf{x}^*$. This proves $\mathbf{x}^*$ strictly maximizes f over U.

The proof of case (2) is the same, except that $D^2f(\mathbf{y})$ is positive definite, making $f(\mathbf{x}^*) < f(\mathbf{x})$ for all $\mathbf{x} \in U$, so $\mathbf{x}^*$ stricly minimizes f over U.

In case (3) $\mathbf{H} = D^2f(\mathbf{x}^*)$ is indefinite. Find unit vectors $\mathbf{u}_i$ with $\mathbf{u}_1^T\mathbf{H}\mathbf{u}_1 < 0$ and $\mathbf{u}_2^T\mathbf{H}\mathbf{u}_2 > 0$. The two mappings, $\mathbf{y} \mapsto \mathbf{u}_i^T D^2f(\mathbf{y})\mathbf{u}_i$ are both continuous in $\mathbf{y}$. Choose $\varepsilon > 0$ small enough that $B_\varepsilon(\mathbf{x}^*) \subset U$ and both values of the quadratic form have the same sign at $\mathbf{y} \in B_\varepsilon(\mathbf{x}^*)$ as they do at $\mathbf{x}^*$.

Then for $\delta < \varepsilon$, let

$$\mathbf{x}_i = \delta\mathbf{u}_i + \mathbf{x}^* \in B_\varepsilon(\mathbf{x}^*) \subset U.$$

It follows that $f(\mathbf{x}_1) < f(\mathbf{x}^*)$ and $f(\mathbf{x}_2) > f(\mathbf{x}^*)$, showing that $\mathbf{x}^*$ is neither a maximum nor a minimum. ∎

17.7 Global Optima and Hessian Minors

We can use the Definite and Semidefinite Matrices Theorems to restate Theorems 17.4.1 and 17.5.1.

Theorem 17.4.1 Restated. *Let* $f\colon U \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *function defined on an subset* $U \subset \mathbb{R}^m$. *Suppose* $\mathbf{x}^*$ *is a critical point of* f *in the interior of* U *and that* U *is star-shaped with respect to* $\mathbf{x}^*$.

1. *If for every* $\mathbf{x} \in U$, *the even order principal minors of the Hessian* $D^2 f(\mathbf{x})$ *are non-negative and the odd order principal minor are non-positive, then* $\mathbf{x}^*$ *maximizes* f *over* U.
2. *If for every* $\mathbf{x} \in U$, *every principal minor of the Hessian* $D^2 f(\mathbf{x})$ *is non-negative, then* $\mathbf{x}^*$ *minimizes* f *over* U.

Theorem 17.5.1 Restated. *Let* $f\colon U \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *function defined on an subset* $U \subset \mathbb{R}^m$. *Suppose* $\mathbf{x}^*$ *is a critical point of* f *in the interior of* U *and that* U *is star-shaped with respect to* $\mathbf{x}^*$.

1. *If for every* $\mathbf{x} \in U$, *the leading principal minors of the Hessian* $\mathbf{H} = D^2 f(\mathbf{x})$ *obey*

$$(-1)^k \det \mathbf{H}_k > 0 \quad \textit{for all } k = 1, \ldots, m$$

then $\mathbf{x}^*$ *strictly maximizes* f *over* U.

2. *If for every* $\mathbf{x} \in U$, *the leading principal minors of the Hessian* $\mathbf{H} = D^2 f(\mathbf{x})$ *obey*

$$\det \mathbf{H}_k > 0 \quad \textit{for all } k = 1, \ldots, m$$

then $\mathbf{x}^*$ *strictly minimizes* f *over* U.

3. *If the Hessian* $D^2 f(\mathbf{x}^*)$ *has leading principal minors that are non-zero and violate the sign patterns in parts (1) and (2), then* $\mathbf{x}^*$ *is neither a maximizer nor minimizer of* f *over* U.

17.8 Second Order Sufficient Conditions at a Critical Point

The next set of second order conditions includes the second order sufficient conditions, points (1) and (2) in the theorem. In these cases, we are able to show that a critical point $\mathbf{x}^*$ is a local max or min.

Theorem 17.8.1. *Let* $f\colon U \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *function defined on an subset* $U \subset \mathbb{R}^m$. *Suppose* $\mathbf{x}^*$ *is a critical point of* f *in the interior of* U.

1. *If the Hessian* $D^2f(\mathbf{x}^*)$ *is negative definite then* $\mathbf{x}^*$ *is a strict local maximizer of* f.
2. *If the Hessian* $D^2f(\mathbf{x}^*)$ *is positive definite then* $\mathbf{x}^*$ *is a strict local minimizer of* f.

Proof. We start with equation (17.3.2):

$$f(\mathbf{x}) - f(\mathbf{x}^*) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}}\left[D^2f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*). \tag{17.3.2}$$

where $\mathbf{y} \in \ell(\mathbf{x}^*, \mathbf{x})$.

By Corollary 16.13.1, the set of positive (negative) definite matrices is an open set. Choose $\varepsilon > 0$ so that $B_\varepsilon(\mathbf{x}^*)$ is contained in that open set. Then choose $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$. The line segment $\ell(\mathbf{x}, \mathbf{x}^*) \in B_\varepsilon(\mathbf{x}^*)$, so $D^2f(\mathbf{y})$ will be positive (negative) definite for $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$.

Then for $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$ and $\mathbf{y} \in \ell(\mathbf{x}, \mathbf{x}^*)$,

$$(\mathbf{x} - \mathbf{x}^*)^{\mathsf{T}}\left[D^2f(\mathbf{y})\right](\mathbf{x} - \mathbf{x}^*) \quad \begin{cases} > 0 & \text{when } D^2f(\mathbf{x}^*) \text{ is positive definite} \\ < 0 & \text{when } D^2f(\mathbf{x}^*) \text{ is negative definite.} \end{cases}$$

Now we apply equation (17.3.2).

It follows that when $D^2f(\mathbf{x}^*)$ is positive definite, for all $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$ with $\mathbf{x} \neq \mathbf{x}^*$, $f(\mathbf{x}) > f(\mathbf{x}^*)$. The point $\mathbf{x}^*$ is a local minimizer, proving item (2).

It also follows that when $D^f(\mathbf{x}^*)$ is negative definite, for all $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$ with $\mathbf{x} \neq \mathbf{x}^*$, $f(\mathbf{x}) < f(\mathbf{x}^*)$. The point $\mathbf{x}^*$ is a local maximizer, proving item (1). ∎

17.9 Second Order Necessary Conditions

Although we will not directly cite Corollary 16.13.2, the fact that the semidefinite matrices are a closed set plays a key role in proving the second order necessary conditions. We use a limiting argument instead of the corollary. That limiting argument could also be adapted to provide an alternative proof to Corollary 16.13.2.

Theorem 17.9.1. *Let $f\colon U \to \mathbb{R}$ be a $\mathcal{C}^2$ function defined on an subset $U \subset \mathbb{R}^m$. Suppose $x^* \in U^0$ is a local max or min. Then x^* is a critical point of f, and:*

1. *If x^* is a local maximizer, the Hessian $D^2f(x^*)$ is negative semidefinite.*
2. *If x^* is a local minimizer, the Hessian $D^2f(x^*)$ is positive semidefinite.*

Proof. That x^* is a critical point of f is Theorem 17.2.2. Since x^* is a critical point, (17.3.2) holds:

$$f(x) - f(x^*) = \frac{1}{2}(x - x^*)^{\mathsf{T}}\left[D^2f(y)\right](x - x^*). \tag{17.3.2}$$

where y is a point in $\ell(x^*, x)$.

If x^* is a local max, we can take $\varepsilon > 0$ so that $f(x) \le f(x^*)$ for $x \in B_\varepsilon(x^*)$. Then

$$0 \ge f(x) - f(x^*) = \frac{1}{2}(x - x^*)^{\mathsf{T}}\left[D^2f(y)\right](x - x^*)$$

for all $x \in B_\varepsilon(x^*)$. Dividing by $\|x - x^*\|^2$, we obtain

$$0 \ge \left(\frac{(x - x^*)}{\|x - x^*\|}\right)^{\mathsf{T}}\left[D^2f(y)\right]\left(\frac{(x - x^*)}{\|x - x^*\|}\right)$$

Let u be any unit vector and define $x_n = x^* + (1/n)u$, so $x_n - x^* = (1/n)u$. For n large, $x_n \in B_\varepsilon(x^*)$ and

$$0 \ge \left(\frac{(x_n - x^*)}{\|x_n - x^*\|}\right)^{\mathsf{T}}\left[D^2f(y_n)\right]\left(\frac{(x_n - x^*)}{\|x_n - x^*\|}\right) \ge u^{\mathsf{T}}\left[D^2f(y_n)\right]u$$

where $y_n \in \ell(x_n, x^*)$. Letting $n \to \infty$, $y_n \to x^*$ and we obtain

$$0 \ge u^{\mathsf{T}}\left[D^2f(x^*)\right]u.$$

since D^2f is continuous. Since this is true for any unit vector, $D^2f(x^*)$ is negative semidefinite.

The case where x^* is a local max is similar. ∎

17.10 Local Optima and Hessian Minors

As with the theorem on global optima, we can use the Definite and Semidefinite Matrices Theorems to restate Theorems 17.8.1 and 17.9.1.

Theorem 17.8.1 Restated. *Let* $f\colon U \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *function defined on an subset* $U \subset \mathbb{R}^m$. *Suppose* $\mathbf{x}^*$ *is a critical point of* f *in the interior of* U.

1. *If the leading principal minors of the Hessian* $\mathbf{H} = D^2 f(\mathbf{x}^*)$ *obey*

$$(-1)^k \det \mathbf{H}_k > 0 \quad \textit{for all } k = 1, \ldots, m$$

   then $\mathbf{x}^*$ *is a strict local maximizer of* f.

2. *If the leading principal minors of the Hessian* $\mathbf{H} = D^2 f(\mathbf{x}^*)$ *obey*

$$\det \mathbf{H}_k > 0 \quad \textit{for all } k = 1, \ldots, m$$

   then $\mathbf{x}^*$ *is a strict local minimizer of* f.

Theorem 17.9.1 Restated. *Let* $f\colon U \to \mathbb{R}$ *be a* $\mathcal{C}^2$ *function defined on an subset* $U \subset \mathbb{R}^m$. *Suppose* $\mathbf{x}^* \in U^0$ *is a local max or min. Then* $\mathbf{x}^*$ *is a critical point of* f, *and:*

1. *If* $\mathbf{x}^*$ *is a local maximizer, the even order principal minors of the Hessian* $D^2 f(\mathbf{x}^*)$ *are non-negative and the odd order principal minors are non-positive.*
2. *If* $\mathbf{x}^*$ *is a local minimizer, every principal minor of the Hessian* $D^2 f(\mathbf{x}^*)$ *is non-negative.*

October 27, 2020