Optimization and Multivariate Calculus

Joe Patten

August 7, 2019

These notes are to accompany Mathematics for Economists by Simon and Blume.

1 Definiteness

1.1 Principal Minors

We can also evaluate the principal minors of A to determine the definiteness of A.

Let A be an $n \times n$ matrix. A kth order principal submatrix is $k \times k$ and is formed by deleting $n-k$ rows, and the same $n-k$ columns. Taking the determinant of a kth order principal submatrix yields a kth order principal minor.

The kth order leading principal submatrix of A, usually written as $|A_k|$, is the left most submatrix in A that is $k \times k$. The determinant of the kth order leading principal submatrix is called the kth order leading principal determinant.

Example

Find all principal minors for the following matrix:

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

First order principal minors:

$$\det(2) = 2 \leftarrow \text{First order leading principal minor}$$
$$\det(2) = 2$$
$$\det(2) = 2$$

Second order principal minors:

$$\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 \leftarrow \text{Second order leading principal minor}$$
$$\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$$
$$\begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3$$

Third order principal minor:

$$\begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 4 \leftarrow \text{Third order leading principal minor}$$

Let A be an $n \times n$ matrix. Then,

- A is **positive definite** iff all of its leading principal minors are positive.
- A is **negative definite** iff leading principal minors alternate in sign, and the 1st order princiapl minor is negative.
- A is **positive semi-definite** iff every principal minor of is is nonnegative.
- A is **negative semi-definite** iff every principal minor of odd order is nonpositive, and every principal minors of even order is nonnegative.
- A is **indefinite** iff A does not have any of these patterns.

Practice

Show that the following matrix is positive definite:

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

2 Unconstrained Optimization

2.1 Optima

Let $f : X \to \mathbb{R}$ where $X \subseteq \mathbb{R}^n$:

Global Optima

- $x^* \in X$ is a **global max** of F on X if $F(x^*) \geq F(x)$ for all $x \in X$
- $x^* \in X$ is a **global min** of F on X if $F(x^*) \leq F(x)$ for all $x \in X$

Strict Global Optima

- $x^* \in X$ is a **strict global max** of F on X if $F(x^*) > F(x)$ for all $x \in X$
- $x^* \in X$ is a **strict global min** of F on X if $F(x^*) < F(x)$ for all $x \in X$

Local Optima

- $x^* \in X$ is a **local max** of F if there is a epsilon-ball $B_\varepsilon(x^*)$ around x^* such that $F(x^*) \geq F(x)$ for all $x \in X$
- $x^* \in X$ is a **local min** of F if there is a epsilon-ball $B_\varepsilon(x^*)$ around x^* such that $F(x^*) \leq F(x)$ for all $x \in X$

Strict Local Optima

- $x^* \in X$ is a **strict local max** of F if there is a epsilon-ball $B_\varepsilon(x^*)$ around x^* such that $F(x^*) > F(x)$ for all $x \in X$
- $x^* \in X$ is a **strict local min** of F if there is a epsilon-ball $B_\varepsilon(x^*)$ around x^* such that $F(x^*) < F(x)$ for all $x \in X$

2.2 First Order Conditions

Before we talk about first order conditions, we need to define what the interior of a set is. Consider the set $X \subseteq \mathbb{R}^n$. X^o is the interior of set X, where X^o is defined as:

$$X^o = \{\in X : \exists B_\varepsilon(x) \subseteq X\}$$

Each element of X^o is an interior point of X.

Theorem: Let $F : X \to \mathbb{R}$ be a C^1 function where $X \subseteq \mathbb{R}^n$. If x^* is a local max or min of F on X and x^* is an interior point of X then:

$$DF_{x^*} = \mathbf{0}$$

Example

Let $F(x,y) = x^3 - y^3 + 9xy$. We can find the "critical points" by setting the first order partial derivatives equal to 0:

$$\frac{\partial F}{\partial x} : 3x^2 + 9y = 0$$
$$\frac{\partial F}{\partial y} : -3y^2 + 9x = 0$$

From the first equation, we find that $y = -\frac{1}{3}x^2$. Substitute this into the second equation:

$$\begin{aligned} 0 &= -3(-\frac{1}{3}x^2)^2 + 9x \\ &= -\frac{1}{3}x^4 + 9x \\ &\Rightarrow x = 0 \text{ or } x = 3 \end{aligned}$$

Plugging these values into either equation gives us the critical points: $(0,0)$ and $(3,-3)$.

Notice that from the theorem above, in order for x^* to be an optimum, it is a necessary condition for all first order partials at x^* to be equal to 0. That being said, having all first order partials equal to 0 does not mean that that point is an optimum. That point is known as a **critical point** and could be either a local max, a local min, or a saddle point. We have to check second order conditions to determine what kind of critical point x^* is.

2.3 Second Order Conditions

Theorem: Let $F : X \to \mathbb{R}$ be a C^2 function where $X \subseteq \mathbb{R}^n$ and X is an open set. Further suppose that x^* is a critical point of F.

- x^* is a strict local max of F if the Hessian, $D^2F_{x^*}$ is negative definite.
- x^* is a strict local min of F if the Hessian, $D^2F_{x^*}$ is positive definite.
- x^* is a saddle point of F (neither a local min or local max) if the Hessian, $D^2F_{x^*}$ is indefinite.

Example

Using the same example as before, $F(x,y) = x^3 - y^3 + 9xy$. The critical points are $(0,0)$ and $(3,-3)$. The Hessian of F is:

$$\begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

At the critical point $(0,0)$, the Hessian is:

$$\begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}$$

Notice that the first order leading principal minor is: $|0| = 0$, and the second order leading principal minor is $\begin{vmatrix} 0 & 9 \\ 9 & 0 \end{vmatrix} = -81$. Notice that the Hessian at $(0,0)$ is indefinite, thus $(0,0)$ is a saddle point.

At the critical point $(3,-3)$, the Hessian is:

$$\begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}$$

Notice that the first order leading principal minor is: $|18| = 18$, and the second order leading principal minor is $\begin{vmatrix} 18 & 9 \\ 9 & 18 \end{vmatrix} = 243$. Notice that the Hessian at $(3, -3)$ is positive definite, thus $(3, -3)$ is a strict local min.

Theorem: Let $F : X \to \mathbb{R}$ be a C^2 function where $X \subseteq \mathbb{R}^n$. Suppose that x^* is an interior point of X and x^* is a local max (respectively min) of F. Then:

1. $DF_{x^*} = \mathbf{0}$
2. $D^2F_{x^*}$ is negative semi-definite (respectively, positive semi-definite)

Theorem: Let $F : X \to \mathbb{R}$ be a C^2 function where $X \subseteq \mathbb{R}^n$ and X is an open, convex set. The following conditions are equivalent (meaning if one condition is true, the other conditions are true):

1. F is a concave function on X
2. $F(y) - F(x) \leq DF_x(y - x) \ \ \forall x, y \in X$
3. $D^2F_{x^*}$ is negative semi-definite $\forall x, y \in X$

Practice

Let $F : X \to \mathbb{R}$ be a C^2 function where $X \subseteq \mathbb{R}^n$ and X is an open, convex set. Show that F is a concave function on $X \Rightarrow$F(y) - F(x) $\leq DF_x(y - x) \ \ \forall x, y \in X$

The following conditions are equivalent:

1. F is a convex function on X
2. $F(y) - F(x) \geq DF_x(y - x) \ \ \forall x, y \in X$
3. $D^2F_{x^*}$ is positive semi-definite $\forall x, y \in X$

Now, assume that F is a concave function on X, then we know that $F(y) - F(x) \leq DF_x(y - x) \ \forall x, y \in X$. Notice that if x^* is a local max or min and in the interior of X, then it follows that $DF_{x^*} = \mathbf{0}$. Thus $F(y) - F(x^*) \leq 0 \Rightarrow F(x^*) \geq F(y) \ \forall y \in X$. Thus, the following follows:

Theorem: If F is a concave function on X and $DF_{x^*} = \mathbf{0}$ for some $x^* \in X$, then x^* is a global max of F on X

Theorem: If F is a convex function on X and $DF_{x^*} = \mathbf{0}$ for some $x^* \in X$, then x^* is a global min of F on X

Exercises

1. Determine the definiteness of the following matrix:

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

2. Find the critical points and classify these as local max, local min, saddle point, or "can't tell":

$$f(x, y, z) = \left(x^2 + 2y^2 + 3z^2\right) e^{-\left(x^2+y^2+z^2\right)}$$