# Optimization and Multivariate Calculus

Joe Patten

August 7, 2019

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

## 1 Multivariate Calculus

### 1.1 Chain Rule

Let $w = f(x, y)$ where $f$ is a differentiable function of $x$ and $y$. Let $x = g(t)$ and $y = h(t)$ where $g$ and $h$ are differentiable functions of $t$. Then by the chain rule:

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

**Example**

Let $w = x^3y^2 - x^2$ and $x = e^t$ and $y = cos(t)$.

$$\begin{aligned}\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}\\ &= \left(3x^2y^2 - 2x\right)\left(e^t\right) + \left(2x^3y\right)\left(-sin(t)\right)\\ &= \left(3e^{2t}cos^2(t) - 2e^t\right)\left(e^t\right) - \left(2e^{3t}cos(t)\right)\left(sin(t)\right)\end{aligned}$$

### 1.2 Total Differential

Recall that when we take a partial derivative, we measure a variable's direct effect on a function (as we keep all other variables constant). If we also want to take into account a variable's indirect effect on a function (i.e. the effect that it has on other variables, which in turn affect the function), then we need to take a total differential.

Consider $z = f(x, y)$. The total differential of $z$ is given by:

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy$$

**Example**

Find the total differential for: $z = 2x\sin(y) - 3x^2y^2$.

$$\begin{aligned}dz &= \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy\\ &= \left(2\sin(y) - 3x^2y^2\right)dx + \left(2x\cos(y) - 6x^2y\right)dy\end{aligned}$$

## 1.3 Implicit Differentiation

Consider the equation $F(x, y) = 0$ where y is defined implicitly as a differentiable function of x. Then,

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}$$

**Example**

Consider $xy^2 + x^3y + 5y - 4 = 0$. Find $\frac{dy}{dx}$:

$$\begin{aligned} \frac{dy}{dx} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \\ &= -\frac{y^2 + 3x^2y}{2xy + x^3 + 5} \\ &= \frac{-y^2 - 3x^2y}{2xy + x^3 + 5} \end{aligned}$$

**Practice**

Use the chain rule to derive the implicit differentiation problem above.

## 1.4 Taylor Series/Polynomial

If $f$ is differentiable of order $n+1$ on interval $I$, then there exists $z$ between points $x$ and $c$, which are on in the interval $I$, such that:

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \frac{f'''(c)}{3!}(x-c)^3 + \ldots + \frac{f^n(c)}{n!}(x-c)^n + R_n(c)$$

where $R_n(c) = \frac{f^{n+1}(z)}{(n+1)!}(x-c)^{n+1}$.

$R_n(c)$ is commonly referred to as the remainder or error. There are many uses of the Taylor polynomial. One use is to approximate the value of a function at a certain point, $x$, given that you know the value of the function at a close point, $c$. The higher the degree of polynomial we use, the closer we will get the the actual value of $f(x)$. You will notice that in each equation below, I have left out the remainder term, thus, we get an approximate value for $f(x)$

First order Taylor polynomial: $f(x) \approx f(c) + f'(c)(x-c)$
Second order Taylor polynomial: $f(x) \approx f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2$
Third order Taylor polynomial: $f(x) \approx f(c) + f'(c)(x-c) + \frac{f''(c)}{2!}(x-c)^2 + \frac{f'''(c)}{3!}(x-c)^3$

**Practice**

Given a function is strictly concave, and $x > c$ (i.e. we are given $f(c)$ and approximating $f(x)$), show that the approximate value for $f(x)$ using a first order Taylor polynomial is greater than the actual value of $f(x)$.

**Note**

A Maclaurin series is a special case of a Taylor series or polynomial. In a Maclaurin series, $c = 0$.

# 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. Consider the function $f(x) = \ln(1+x)$.
   (a) Calculate $f(.5)$.
   (b) Using a first order Taylor polynomial, approximate $f(.5)$ using $x_0 = 0$.
   (c) Using a second order Taylor polynomial, approximate $f(.5)$ using $x_0 = 0$.
   (d) Using a third order Taylor polynomial, approximate $f(.5)$ using $x_0 = 0$.

2. Differentiate implicitly to find $\frac{dy}{dx}$:

$$x^2 - 3xy + y^2 - 2x + y - 5 = 0$$

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

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