Multivariable Calculus – Lecture #8 Notes

In this lecture we continue the discussion of extrema of functions of several variables. In particular, we'll discuss a protocol for finding extrema in a bounded region, we'll look at some applications in economics, and we'll address how to handle multiple constraints. We'll also get started on the idea of integration of a function.

Recall from the previous lecture:
Extreme Value Theorem – Suppose a real-valued function $f(x_1,\cdots,x_n)$ is continuous on a closed and bounded domain $D\subset\mathbf{R}^n$. Then this function must attain its maximum and minimum values somewhere within this domain. [Note: You can think of a closed set as one for which any convergent sequence in this set has its limit in the set. In particular, this means that the boundary of the set must be included.]

For a function $f(x_1,\cdots,x_n)$ of several variables, extrema will occur at:

(a) Stationary points (points where $\overrightarrow{\nabla f}=\mathbf{0}$, i.e. where all of the partial derivatives of f vanish); or
(b) Non-differentiable points (points where the given function is not differentiable); or
(c) On the boundary of the domain.

The stationary points and non-differentiable points are ***critical points***. The analysis of a typical problem usually is divided into two steps: (1) seeking critical points in the interior of the region (though they could occur on the boundary), and (2) seeking additional candidates for extrema on the boundary of the region using the Method of Lagrange Multipliers.

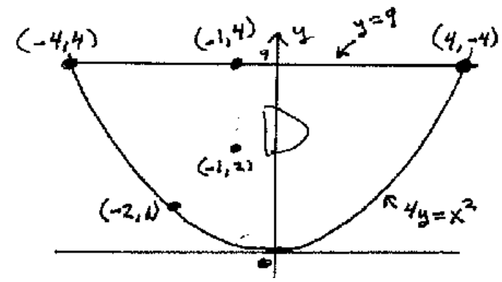
$$\begin{Bmatrix} \overrightarrow{\nabla f}=\lambda\overrightarrow{\nabla g} \\ g(x_1,\cdots,x_n)=c \end{Bmatrix}$$

Associated with this theorem is the following protocol for identifying extreme of functions in a bounded region:

(1) First **seek stationary points in the interior** of the bounded domain (which might possibly occur on the boundary of the domain) as well as other critical points, e.g. points of non-differentiability; then
(2) **Use the Method of Lagrange Multipliers to identify candidates for extrema on the boundary** of the domain (assuming the boundary is described by level sets of a differentiable function). If the boundary consists of several distinct pieces, you will also have to include where those pieces intersect as possible locations for extrema.

It must also be said that we often wish to find extrema of functions in unbounded domains. In this case we proceed as above, but determining which points yield maxima or minima may require some additional argument. Often, these methods come down to identifying a short list of candidates for extrema and evaluation of the given function to determine which yield relative and absolute maxima and minima.

Example: Find the maximum and minimum values of the function $f(x,y)=x^2+y^2+2x-4y+5$ in the region D bounded by the curves $4y=x^2$ and $y=4$.



Solution: (1) We first seek interior critical points: These occur where $\begin{Bmatrix} f_x=2x+2=0 \\ f_y=2y-4=0 \end{Bmatrix} \Rightarrow (x,y)=(-1,2)$. You can apply the 2nd Derivative Test, if you wish, to see that this will yield a local minimum.

(2) We next examine the boundary for possible extrema. Along the top edge ($y=4$) we can just substitute to get $f=x^2+2x+5$ and basic Calculus methods give a critical point where $2x+2=0$, i.e. $x=-1$. So the point $(-1,4)$ is a candidate for a maximum or a minimum. For the other boundary ($4y=x^2$) we could also substitute, but if we use the Method of

 Revised January 7, 2017

Lagrange Multipliers we first rewrite this constraint as $g(x,y) = 4y - x^2 = 0$. The optimality condition gives:

$$\begin{Bmatrix} 2x+2 = \lambda \cdot (-2x) \\ 2y-4 = \lambda \cdot (4) \end{Bmatrix} \Rightarrow \frac{2x+2}{2y-4} = \frac{-2x}{4} \Rightarrow 8x+8 = -4xy+8x \Rightarrow xy = -2$$

Combining this with the constraint $4y = x^2$ we get $x^3 = -8$, so $x = -2$ and $y = 1$, i.e. the point $(-2,1)$.

The boundaries meet at the points $(-4,4)$ and $(4,4)$ and these must also be considered as candidates for possible extrema. Evaluation and comparison gives, as we start with the stationary point and then work our way sequentially around the boundary:

Candidate point	**Value**	**Notes**
$(-1,2)$	$f(-1,2) = \boxed{0}$	absolute **minimum** at this interior stationary point
$(-4,4)$	$f(-4,4) = \boxed{13}$	relative maximum at this boundary point
$(-1,4)$	$f(-1,4) = \boxed{4}$	neither maximum nor minimum
$(4,4)$	$f(4,4) = \boxed{29}$	absolute **maximum**
$(-2,1)$	$f(-2,1) = \boxed{2}$	neither maximum nor minimum

Lagrange Multiplier methods in Economics

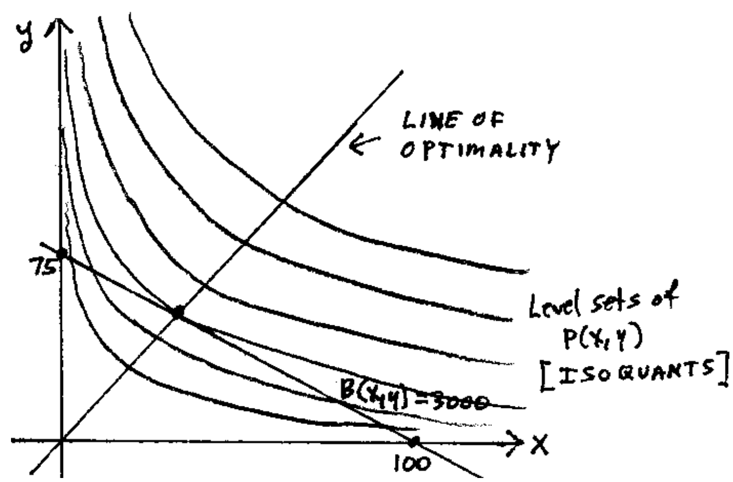
Perhaps the most common setting where constrained optimization is encountered is in the field of Economics. Indeed, you could almost define a significant portion of economics by this. Here's a typical problem that you might encounter:

Problem: Suppose production at a factory is modeled by $P(x,y) = 400x^{1/2}y^{3/4}$ where x represents the number of units of labor and y represents the number of units of capital. [A production model of this form ($P = ax^m y^n$) is known as a *Cobb-Douglas* model.] If labor costs \$30/unit and capital costs \$40/unit, find the optimal combination of labor and capital that yields the greatest production given a budget of \$3000. [Note: These figures might correspond to hourly production.]

Solution: We want to maximize $P(x,y) = 400x^{1/2}y^{3/4}$ subject to the constraint $B(x,y) = 30x + 40y = 3000$. Extrema will occur where:

$$\overrightarrow{\nabla P} = \lambda \overrightarrow{\nabla B} \Rightarrow \begin{Bmatrix} P_x = \lambda B_x \\ P_y = \lambda B_y \end{Bmatrix} \Rightarrow \begin{Bmatrix} 200x^{-1/2}y^{3/4} = \lambda \cdot 30 \\ 300x^{1/2}y^{-1/4} = \lambda \cdot 40 \end{Bmatrix} \Rightarrow \frac{2y}{3x} = \frac{3}{4} \Rightarrow \boxed{y = \tfrac{9}{8}x}$$

Generally, we see that under optimal conditions $\dfrac{P_x}{P_y} = \dfrac{B_x}{B_y} = \dfrac{\text{unit price of labor}}{\text{unit price of capital}}$. That is, under optimal conditions the ratio of the marginal productivities is equal to the ration of the unit prices.



In the specific problem, the equation $y = \frac{9}{8}x$ represents a "line of optimality" independent of whatever budget we have with which to work. If we solve this simultaneously with the given budget constraint we get that $30x + 40\left(\frac{9}{8}x\right) = 75x = 3000$, so $x = 40$ units of labor and $y = \frac{9}{8}x = 45$ units of capital. Under these optimal conditions, the maximum productivity will be $P_{\max} = P(40,45) = 400(40)^{1/2}(45)^{3/4} \cong 43954.098$. There's a reason why we're using an excessive number of significant figures. To see why, let's consider what would happen if we were to marginally increase the budget.

If we instead had \$3001 to spend optimally, how would things change? Optimally we would still have to maintain the condition that $y = \frac{9}{8}x$, so the only change would be that $75x = 3001$ which would give that $x = 40.1333\ldots$ units of labor and $y = 45.015$ units of capital. Using these, the new maximum production would be $P_{\max} = P(40.0133\ldots, 45.015) \cong 43972.413$. So get an increase in production of $\Delta P \cong 18.314$. If we think of the budget B as a parameter that can be adjusted, we see that by increasing the budget by \$1 we have that $\frac{\Delta P}{\Delta B} \cong 18.314$. Curiously, we can calculate that the value of the Lagrange Multiplier is $\lambda = \frac{200(40)^{-1/2}(45)^{3/4}}{30} \cong 18.3142$. This seems unlikely to have occurred by chance, so what's the explanation? Everything can be explained using the Chain Rule.

The method of solution can be thought of sequentially as:

$$B \xrightarrow{\text{solve via L.M.}} (x(B), y(B)) \to P(x(B), y(B)) = P_{\max}$$

The Chain Rule plus substitution of the Lagrange Multiplier conditions gives that:

$$\frac{dP}{dB} = \frac{\partial P}{\partial x}\frac{\partial x}{\partial B} + \frac{\partial P}{\partial y}\frac{\partial y}{\partial B} = \left(\lambda \frac{\partial B}{\partial x}\right)\frac{\partial x}{\partial B} + \left(\lambda \frac{\partial B}{\partial y}\right)\frac{\partial y}{\partial B} = \lambda\left(\frac{\partial B}{\partial x}\frac{\partial x}{\partial B} + \frac{\partial B}{\partial y}\frac{\partial y}{\partial B}\right).$$

To make sense of the somewhat cryptic expression inside the parentheses (where B appears both as a parameter and as a function of x and y, consider the somewhat silly sequence:

$$B \xrightarrow{\text{solve via L.M.}} (x(B), y(B)) \to B(x(B), y(B)) = B$$

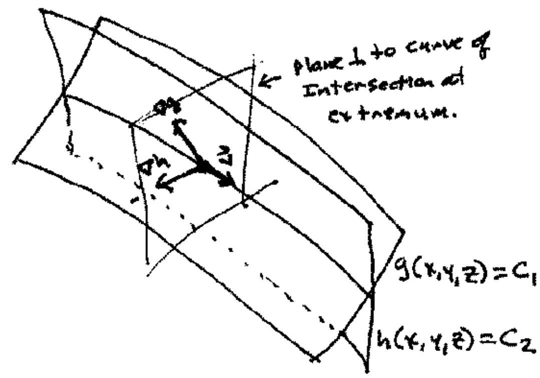
This basically says that if you are given B dollars to spend optimally and someone then asks you how much money you must then spend, the answer is simply "as much money as I have to work with". The Chain Rules then gives that $\frac{dB}{dB} = \frac{\partial B}{\partial x}\frac{\partial x}{\partial B} + \frac{\partial B}{\partial y}\frac{\partial y}{\partial B} = 1$, so $\frac{dP}{dB} = \lambda\left(\frac{\partial B}{\partial x}\frac{\partial x}{\partial B} + \frac{\partial B}{\partial y}\frac{\partial y}{\partial B}\right) = \lambda \cdot 1 = \lambda$. This explains why we found that for a small change in budget we have $\frac{\Delta P}{\Delta B} \cong \lambda$.

Speaking colloquially, we might say that λ measures "how much more bang for the buck".

Reciprocity: In a problem like the previous one, the Lagrange Multiplier condition $\overrightarrow{\nabla P} = \lambda \overrightarrow{\nabla B}$ simply describes the fact that under optimal conditions $\overrightarrow{\nabla P}$ and $\overrightarrow{\nabla B}$ will be parallel.. It could just as easily have been expressed as $\overrightarrow{\nabla B} = \frac{1}{\lambda}\overrightarrow{\nabla P}$. This can be interpreted as saying that the problem of maximizing production subject to a fixed budget is fundamentally the same as minimizing cost subject to a fixed production. Essentially, efficiency is the real point and we can interchange the roles of what quantity is being optimized and what quantity is being constrained.

The Method of Lagrange Multipliers with Multiple Constraints

Suppose we wish to find the extrema of a function $f(x,y,z)$ subject to two constraints: $g(x,y,z)=c_1$ and $h(x,y,z)=c_2$. Let's further assume that the surfaces represent by both constraints are smooth surfaces and that they intersect in a smooth curve, i.e. a curve that can be parameterized with differentiable component functions. It is worth noting that in a poorly-posed problem the constraints might not even be compatible. Even more subtle, the two constraints might only intersect tangentially, so we will further assume that they have a "clean intersection" or a "transverse intersection" in which the normal vectors to these constraint surface are never parallel along their intersection.



Under these conditions, if $\mathbf{r}(t)=\langle x(t),y(t),z(t)\rangle$ parameterizes the intersection curve, then at any extremum the function $f(\mathbf{r}(t))=f(x(t),y(t),z(t))$ would have a stationary point. By the Chain Rule, this means that $\frac{d}{dt}\left[f(\mathbf{r}(t))\right]=\overrightarrow{\nabla f}\cdot\mathbf{v}=0$ where **v** is the velocity (tangent) vector to the intersection curve. This means that $\overrightarrow{\nabla f}\perp\mathbf{v}$ at an extremum. But we know that $\overrightarrow{\nabla g}$ and $\overrightarrow{\nabla h}$ are perpendicular at this point to the respective constraint surfaces. As long as we know that $\overrightarrow{\nabla g}$ and $\overrightarrow{\nabla h}$ are not parallel, they will span the plane perpendicular to the vector **v**, so we must therefore be able to express $\overrightarrow{\nabla f}=\lambda_1\overrightarrow{\nabla g}+\lambda_2\overrightarrow{\nabla h}$ for some scalars (Lagrange Multipliers) λ_1 and λ_2. Together with the two constraints, this is our enhanced Method of Lagrange Multipliers for multiple constraints: $\left\{\begin{matrix}\overrightarrow{\nabla f}=\lambda_1\overrightarrow{\nabla g}+\lambda_2\overrightarrow{\nabla h}\\ g(x,y,z)=c_1\\ h(x,y,z)=c_2\end{matrix}\right\}$. Note that this yields a total of 5 equations in the 5 unknowns $\{x,y,z,\lambda_1,\lambda_2\}$. It may not be easy, but these can in principle be solved.

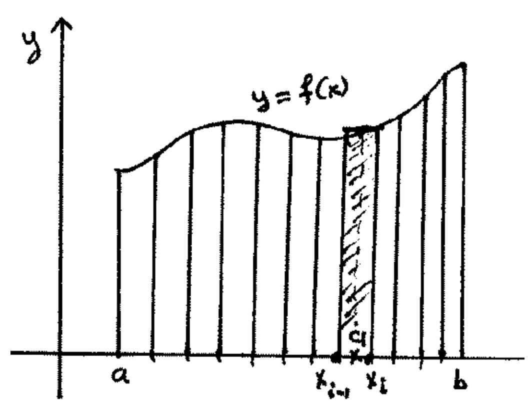
This method can, under ideal conditions, be extended to a function of n variables with m constraints, i.e. to find the extrema of $f(x_1,\cdots,x_n)$ subject to the constraints $\left\{\begin{matrix}g_1(x_1,\cdots,x_n)=c_1\\ \vdots\\ g_m(x_1,\cdots,x_n)=c_m\end{matrix}\right\}$ we would solve the equations:

$$\left\{\begin{matrix}\overrightarrow{\nabla f}=\lambda_1\overrightarrow{\nabla g_1}+\cdots+\lambda_m\overrightarrow{\nabla g_m}\\ g_1(x_1,\cdots,x_n)=c_1\\ \vdots\\ g_m(x_1,\cdots,x_n)=c_m\end{matrix}\right\}$$

($n+m$ equations in the $n+m$ unknowns $\{x_1,\cdots,x_n,\lambda_1,\cdots,\lambda_m\}$.

Integration of functions of several variables over regions in $\mathbf{R}^2$

Integration is really about measurement – specifically the measurement of an aggregate amount of *something*. Recall the ideas behind integration of a function of one variable over an interval, i.e. $\int_a^b f(x)dx$, the definite integral. The motivating example was most likely finding the area of a region under the graph and above the horizontal axis of a positive function $f(x)$ defined on an interval $[a,b]$. The method utilized to calculate this was the **Method of Riemann Sums**. This consists of four steps:



1) ***Partition the domain***: Choose additional points $\{x_i\}$ so that $a = x_0 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$ and let $\Delta x_i = x_i - x_{i-1}$ be the width of the *i*-th interval.

2) ***Approximate what is to be measured***: Choose a sample point c_i in each interval, and approximate the area of the *i*-th rectangle as $\Delta A_i \cong f(c_i)\Delta x_i$.

3) ***Sum***: Add up these approximate areas to get an approximation of the total area, i.e. $A = \sum_{i=1}^{n} \Delta A_i \cong \sum_{i=1}^{n} f(c_i)\Delta x_i$.

4) ***Limit***: Refine the partition by inserting points so that the width of the largest interval gets progressively smaller and find the limit of the approximate total area (if it exists) as these widths become arbitrarily small. That is, if we denote the ***mesh of the partition*** by $|\Delta| = \max_{1\le i\le n}(\Delta x_i)$, we define $\lim_{|\Delta|\to 0}\left(\sum_{i=1}^{n} f(c_i)\Delta x_i\right) = \int_a^b f(x)dx$ if this limit exists independent of any choices.

If this limit exists we say that the function $f(x)$ is (Riemann) **integrable** over the interval $[a,b]$. It is proven (or should be proven) in single-variable Calculus that if $f(x)$ is continuous or piecewise-continuous on the interval $[a,b]$ with only jump discontinuities, then $f(x)$ will be integrable.

It must be noted that the calculation of definite integrals in one-variable Calculus is made much simpler via the application of the ***Fundamental Theorem of Calculus***.

Once the idea of the definite integral is established, it is simple to observe that there's really no reason to assume $f(x)$ to be positive on the interval $[a,b]$. The definition of $\int_a^b f(x)dx$ still makes sense, but must be interpreted then as a "signed area" where regions above the horizontal axis are counted as positive area and regions below the horizontal axis are counted as negative area. The integral then measures the sum of these values.

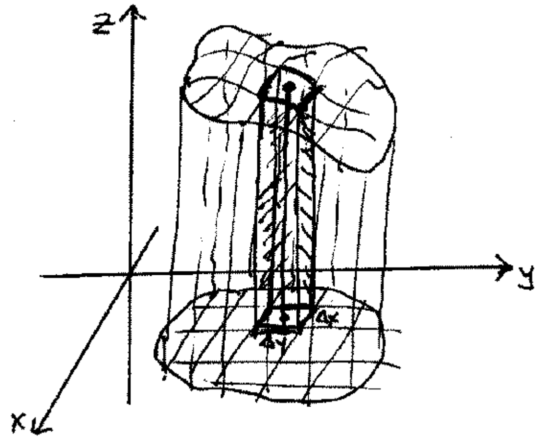
More significantly, the definite integral doesn't have to be interpreted or constructed as an area at all. Here are two other simple examples of some significance:

(a) Suppose a bar of material of variable density is located in such a way that it occupies the interval $[a,b]$. If $\sigma(x)$ measures the density at each point in the interval (perhaps in units such as grams per centimeter), the same Method of Riemann Sums would give $\Delta m_i \cong \sigma(c_i)\Delta x_i$ for the approximate mass of the *i*-th interval and $\int_a^b \sigma(x)dx$ would then give the total mass of this bar of material. If the density measures electric charge density, then the integral will measure the total charge. Indeed, if $\sigma(x)$ measures the density of any quantity whatsoever in this interval, the integral $\int_a^b \sigma(x)dx$ will measure the <u>total amount</u> of this quantity.

(b) Suppose that over the time interval $a \le t \le b$ an object moves spatially (perhaps along a straight line) in such a way that its velocity at any time *t* is given by a function $v(t)$. Over a small time interval, this object would be displaced by an amount $\Delta s_i \cong v(t_i)\Delta t_i$, i.e. rate times time gives distance, and the total displacement (including both positive and negative displacements, depending upon when the velocity is positive or negative) would be approximately $\sum_{i=1}^{n} v(t_i)\Delta t_i$, so in the limit we interpret $\int_a^b v(t)dt$ as the total displacement of this moving object.

How might we think about the idea of the integral of a function of two (or more) variables? We can proceed by analogy – at least to get started.

Suppose $f(x,y)$ is a function of two variables defined over a (closed and bounded) domain *D*. If this function has positive values in this domain, we can consider its graph $z=f(x,y)$ over this domain and, by analogy, try to find a way of measuring the ***volume under this graph*** (and above the *xy*-plane and within the vertical "curtain wall" lying above the boundary of the domain *D*). Proceeding as before using the Method of Riemann Sums, we would:



1) ***Partition the domain***: This time we would have to chop up the two-dimensional domain *D* into uniformly small pieces D_i – perhaps mainly small rectangular pieces, but not necessarily. Let ΔA_i denote the area of the *i*-th piece and define the ***mesh of the partition*** by $|\Delta| = \max\limits_{1\le i\le n}\left(\text{diameter}(D_i)\right)$.

2) ***Approximate what is to be measured***: Choose a sample point (x_i, y_i) in each piece D_i, and approximate the volume of the vertical shaft above this piece by $\Delta V_i \cong f(x_i,y_i)\Delta A_i$.

3) ***Sum***: Add up these approximate volumes to get an approximation of the total volume, i.e.
$V=\sum_{i=1}^{n}\Delta V_i \cong \sum_{i=1}^{n} f(x_i,y_i)\Delta A_i$.

4) ***Limit***: Refine the partition in such a way that the ***mesh of the partition*** $|\Delta| = \max\limits_{1\le i\le n}\left(\text{diameter}(D_i)\right)$ becomes arbitrarily small. We then define $\lim\limits_{|\Delta|\to 0}\left(\sum_{i=1}^{n} f(x_i,y_i)\Delta A_i\right)=\iint_D f(x,y)\,dA$ if this limit exists independent of any choices. We would then say that this function is integrable over this domain and we refer to $\iint_D f(x,y)\,dA$ as the (double) integral of the function $f(x,y)$ over the domain *D*.

The procedure above is perfectly doable for a computer but, as was the case in single variable Calculus when we used the Fundamental Theorem of Calculus as an alternate way to calculate integrals, we will want to discover simple ways to calculate such a multiple integral (when possible) without the need of a computer.

Once again, it's not really necessary to assume that $f(x,y)$ is positive everywhere in the domain *D*. The definition of $\iint_D f(x,y)\,dA$ still makes sense, but must be interpreted then as a "signed volume" where regions above the horizontal *xy*-plane are counted as positive volume and regions below the horizontal *xy*-plane are counted as negative volume. The integral then measures the sum of these values.

The integral of a function of two variables over a domain *D* doesn't have to measure volume or signed volume. Suppose that $\sigma(x,y)$ measures the density of an object that occupies a domain *D* in the *xy*-plane (perhaps measured in units of grams per square centimeter). If we were to cut up the domain *D* into uniformly small pieces D_i, we might then approximate the mass of each small piece as $\Delta m_i \cong \sigma(x_i,y_i)\Delta A_i$ and proceed as before. We would then interpret the integral as the total mass of this object (often referred to as a *lamina*), i.e. $\boxed{\text{mass}(D)=\iint_D \sigma(x,y)\,dA}$. If the density measures electrical charge, the integral $\iint_D \sigma(x,y)\,dA$ will then give the total charge (summing both positive and negatively charged regions to produce the net charge). If $\sigma(x, y)$ measures population density in a region D, then $\iint_D \sigma(x, y)\,dA$ will give the total population in the region.

Perhaps the simplest measurement we might want to consider would be just the total area of a region D. In this case, this involves only summing up the areas of individual pieces, and in the limit we get simply that $\boxed{\text{Area}(D) = \iint_D dA}$.

In the next lecture, we'll describe a range of additional applications of multiple integrals and develop techniques for calculating them. We'll also get started on triple integrals.

Notes by Robert Winters and Renée Chipman