March 13, 2019

LECTURE 13: OPTIMIZATION.

110.211 HONORS MULTIVARIABLE CALCULUS
PROFESSOR RICHARD BROWN

Synopsis. Just as in single variable calculus, optimizing a function of one variable is a matter for the Extreme Value Theorem and local extrema. But often, situations arise where the objective function involves more than one variable. In this case, there are usually relationships between the variables that allow for rewriting the function as a function of one variable. This is a form of constrained optimization that generalizes well to multivariable calculus. Today we explore this idea, using geometry to "see" our way through to a technique. This leads to the technique of Lagrange multipliers, which we develop here.

Helpful Documents. Mathematica: `LagrangeMult`.

One variable optimization. Recall optimization in single variable calculus:

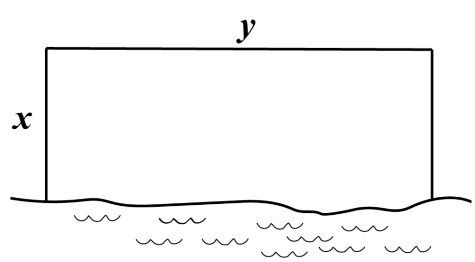

FIGURE 13.1. A fenced yeard along a river.

Example 13.1. Using 1800 linear feet of fencing, construct a rectangular yard along a straight river with the largest area possible. The idea here is to maximize area of a rectangular region. Given the two unknowns of length and width, say, x and y, maximize area $A = xy$. Of course, there is a constraint in that you can only use up to 1800 feet of fencing. Mathematically speaking, this means that $1800 = 2x + y$, given the arrangement of the rectangle in Figure 13.1. We call the area equation here the *objective function*, and the perimater fencing equation the *constraint*.

The constraint facilitates calculation by

- allowing us to change the objective function, via substitution, into a function of only one variable, and
- allows us to use single variable calculus techniques to help locate the extrema of the objective function within the constraints.

Now, since $1800 = 2x + y$, we know $y = 1800 - 2x$, so that

$$A = xy = x(1800 - 2x) = 1800x - 2x^2.$$

This is a clue that we are on the right track here, as $A(x)$ has a graph which is a parabola opening down (the leading coefficient is negative). Hence it will have a max at the vertex. We also know that the variables must be nonnegative numbers, as they denote lengths. Hence $0 \leq x \leq 900$ and $0 \leq y \leq 1800$. Hence $A(x)$ has a domain $[0, 900]$ and by the Extreme Value Theorem, must achieve its maximum either at an endpoint or at a critical point. And as $A(x)$ is differentiable, all critical points will occur at places where $a'(x) = 0$. Here

$$A'(x) = 1800 - 4x = 0 \quad \text{is solved only by} \quad x = 450.$$

Immediately, using the Second Derivative Test, we also see that $A''(450) = -4 > 0$. Hence $x = 450$ corresponds to a local maximum, and since $A(0) = A(900) = 0$, $x = 450$ corresponds to a global maximum.

The solution, then, is to construct the pen in Figure 13.1 with $x = 450$ feet, and $y = 900$ feet.

Here is a different viewpoint of the same problem: Leave the function $A = A(x, y) = xy$ as a function on two variables, and consider the level sets of $A(x, y)$ on the domain

$$R = \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; 0 \leq x \leq 900, 0 \leq y \leq 1800\right\},$$

a few of which are graphed in Figure 13.2. Also on R an in the figure, we can graph the constraint curve (as the red line), thinking of it as the 0-level set of the function $P(x, y) = 2x + y - 1800$. Now, if we are forced to stay on the constraint line, is there a place on this red line where we touch or cross the level set of $A(x, y)$ corresponding to the largest area? One can possibly see it in the figure. But can one "calculate" it?

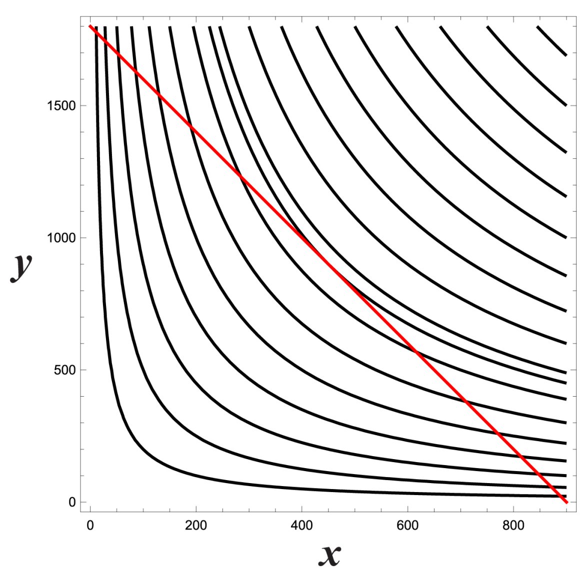

FIGURE 13.2. Level set of $A(x, y) = xy$ in black, and 0-level set of $P(x, y) = 1800 - 2x - y$ in red.

In this new approach, both the objective function and the constraint are left as functions of the two variables. And we search for a geometric solution to locating an extremum of one function constrained by a second one. One can see in the figure that, as we move along the red line, we are cutting through the level sets of A for a while. At some point, we go tangent to a particular level set and then we start cutting through level sets of A again, although in the other direction (first form lower to higher, then from higher to lower values of A). SO what to we see as the values of A along the red line? We see A rising for a while (cutting through blue lines of increasing A, topping out at some point (the red line becomes tangent to a blue line), then declining in value (again cutting through level set of decreasing A). We have found our maximum of A along the red line via the point of tangency with a blue line!

This new, geometric, idea of optimization can be generalized: Optimize $f : X \subset \mathbb{R}^n \to \mathbb{R}$, subject to $g : X \to \mathbb{R}$, where $g(\mathbf{x}) = c$. We look for extrema of f while constrained to the c-level set of g. We note here that the idea we started with, that of using the constraint to remove one of the variables in the objective function is less helpful in this multivariable setting. And possibly impossible:

Example 13.2. Maximize $f(x, y, z) = x^2 + 3y^2 + y^2z^4$, subject to $g(x, y, z) = e^{xy} - xyz + \cos\left(\frac{xy}{z}\right) = 2$. Try to solve for one of the variables in g as a function of the other two, and substitute that into f to remove a variable!

Theorem 13.3. *For $X \subset \mathbb{R}^n$ open, $f, g : X \to \mathbb{R}$ both C^1-functions, let*

$$S = \left\{ \mathbf{x} \in X \;\middle|\; g(\mathbf{x}) = c \right\}$$

be the c-level set of g. Then, if $f\big|_S$ has an extremum at $\mathbf{x}_0 \in S$, where $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$, then $\exists \lambda \in \mathbb{R}$ such that

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0).$$

Some notes:

- The extrema of f will happen at places where ∇f is a multiple of ∇g, as vectors. These wind up being places of tangency between level sets, and places where often the level set of g stops cutting through the level sets of f (for a moment).
- The equation $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$ is actually set of n equaitons (nonlinear) in $n+1$ unknowns (each component of the vector $\mathbf{x}$, along with the real number λ). So there are lots of solutions!
- But if we add in the constraint itself, we arrive at $n+1$ equations in $n+1$ unknowns:

$$\begin{aligned} f_{x_1}(\mathbf{x}) &= \lambda g_{x_1}(\mathbf{x}) \\ &\vdots \\ f_{x_n}(\mathbf{x}) &= \lambda g_{x_n}(\mathbf{x}) \\ g(\mathbf{x}) &= c. \end{aligned}$$

- The variable λ is called a *Lagrange multiplier*. It's actual value is not nearly as important as its existence!

Example 13.4. Identify all critical points of $f(x, y) = 5x + 2y$, subject to $g(x, y) = 5x^2 + 2y^2 = 14$.

Here, $\nabla f(\mathbf{x}) = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $\nabla g(\mathbf{x}) = \begin{bmatrix} 10x \\ 4y \end{bmatrix}$. The system is then

$$\begin{aligned} f_x(\mathbf{x}) &= \lambda g_x(\mathbf{x}) & \qquad 5 &= \lambda 10x \\ f_y(\mathbf{x}) &= \lambda g_y(\mathbf{x}) & 2 &= \lambda 4y \\ g(\mathbf{x}) &= c & 5x^2 + 2y^2 &= 14. \end{aligned}$$

Solving, we find by the first and second equations, that $x = \frac{1}{2\lambda} = y$. So the last equations becomes

$$\frac{5}{4\lambda^2} + \frac{2}{4\lambda^2} = 14, \quad \Longrightarrow \lambda = \pm \frac{1}{2\sqrt{2}}.$$

Hence the critical points are $(x, y) = \left(\sqrt{2}, \sqrt{2}\right)$ and $(x, y) = \left(-\sqrt{2}, -\sqrt{2}\right)$

Geometrically, one can see whether these points are extrema or not, and why the gradient condition is quite telling. And analytically?

Handling multiple constraints is done in the same general fashion:

- Each constraint tends to reduce the number of independent variables by one.
- Each constraint tends to reduce the dimension of the space that we evaluate the objective function along by one.

- in $\mathbb{R}^3$, one objective function has level sets which are generically surfaces. Each constraint will also have level sets which are mostly surfaces. two surfaces typically meet in a curve. We then evaluate the objective function along this curve, looking for extrema in a very single variable calculus fashion.

Example 13.5. Find the extrema of $f(x,y,z) = 2x + y^2 - z^2$, subject to $g_1(x,y,z) = x - 2y = 0$ and $g_2(x,y,z) = x + z = 0$.

Here, one could simply replace z with $-x$ and y with $\frac{x}{2}$, and look for extrema of $f(x) = 2x + \frac{x^2}{4} - x^2 = 2x - \frac{3}{4}x^2$. One would find that $x = \frac{4}{3}$ is the only extremum and that it is a maximum. So the point $\mathbf{x}_0 = \left(\frac{4}{3}, \frac{2}{3}, -\frac{4}{3}\right)$ is the only critical point of f.

Geometrically, How do we construct a system that we can solve for?

Theorem 13.6. *For $X \subset \mathbb{R}^n$ open, $f, g_1, \ldots, g_k : X \to \mathbb{R}$ be C^1-functions, with $k < n$. Let*

$$S = \left\{\mathbf{x} \in X \;\middle|\; g_1(\mathbf{x}) = c_1, \ldots, g_k(\mathbf{x}) = c_k\right\}$$

be the intersection of the level sets of the g_i, $i = 1, \ldots, k$. Then, if $f\big|_S$ has an extremum at $\mathbf{x}_0 \in S$, where $\nabla g_1(\mathbf{x}_0), \ldots, g_k(\mathbf{x}_0)$ are all linearly independent as vectors, then there exist scalars $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ such that

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g_1(\mathbf{x}_0) + \ldots + \lambda_k \nabla g_k(\mathbf{x}_0). \tag{13.1}$$

Notes:

- Recall that linear independence as vectors as means that every vector must be nonzero!
- Basically, as a vector and at an extremum, $\nabla f(\mathbf{x}_0)$ must be in the space spanned by the $\nabla g_i(\mathbf{x}_0)$, for $i = 1, \ldots, k$.

Example 13.7. In Example 13.5 above, we sought the extrema of $f(x,y,z) = 2x + y^2 - z^2$, subject to the two constraints $g_1(x,y,z) = x - 2y = 0$ and $g_2(x,y,z) = x + z = 0$. To use Theorem 13.6, we form Equation 13.1 directly via the vectors

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2 \\ 2y \\ -2z \end{bmatrix}, \quad \nabla g_1(\mathbf{x}) = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \text{and} \quad \nabla g_2(\mathbf{x}) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Here, the constraint vectors are linearly independent everywhere (why is this?), so the system is

$$\begin{aligned}
f_x(\mathbf{x}) &= \lambda_1 \frac{\partial g_1}{\partial x}(\mathbf{x}) + \lambda_2 \frac{\partial g_2}{\partial x}(\mathbf{x}) & 2 &= \lambda_1 + \lambda_2 \\
f_y(\mathbf{x}) &= \lambda_1 \frac{\partial g_1}{\partial y}(\mathbf{x}) + \lambda_2 \frac{\partial g_2}{\partial y}(\mathbf{x}) & 2y &= -2\lambda_1 \\
f_z(\mathbf{x}) &= \lambda_1 \frac{\partial g_1}{\partial z}(\mathbf{x}) + \lambda_2 \frac{\partial g_2}{\partial z}(\mathbf{x}) & -2z &= \lambda_2 \\
g_1(\mathbf{x}) &= c_1 & x - 2y &= 0 \\
g_2(\mathbf{x}) &= c_2 & x + z &= 0.
\end{aligned}$$

There are many ways to solve these 5 equations in 5 unknowns. One way it so eliminate the lambdas in the first equation via substitution using the second and third. One obtains

$\lambda_1 = -y$ and $\lambda_2 = -2z$, so that the first equation is $2 = -y - 2z$. And eliminating x in the last two equations yields the single equation $0 = 2y + z$. Together, the system

$$\begin{aligned} 2 &= -y - 2z \\ 0 &= 2y + z \end{aligned}$$

is solved by $z = -\frac{4}{3}$ and $y = \frac{2}{3}$. One then calculates $x = \frac{4}{3}$, so that the only critical point of f is again $\mathbf{x}_0 = \left(\frac{4}{3}, \frac{2}{3}, -\frac{4}{3}\right)$.