Lecture 20: Convexity and Optimization

We say that if f is a once continuously differentiable function on an interval I, and x is a point in the interior of I that x is a *critical point* of f if

$$f'(x) = 0.$$

Critical points of once continuously differentiable functions are important because they are the only points that can be local maxima or minima.

In the context of critical points, the second derivative of a function f is important because it helps in determining whether a critical point is a local maximum of minimum.

First derivative test Let f be a once continuously differentiable function on an interval I and let x be a critical point. Suppose there is some open interval (a, b) containing x so that for $y \in (a, b)$ with $y < x$, we have $f'(x) > 0$ and so that for $y \in (a, b)$ with $y > x$, we have $f'(x) < 0$, then f has a local maximum at x. If on the other hand, we have $f'(y) < 0$ when $y \in (a, b)$ and $y < x$ and $f'(y) < 0$ when $y \in (a, b)$ and $y > x$ then f has a local minimum at x.

Proof of First derivative test We prove the local maximum case. Suppose $y \in (a, b)$ with $y \neq x$ and $f(y) \geq f(x)$. If $y < x$, by the mean value theorem, there is some c with $y < c < x$ so that

$$\frac{f(x) - f(y)}{x - y} = f'(c).$$

We know, by assumption, that $f'(c) > 0$. Thus $f(x) > f(y)$ which is a contradiction. Suppose instead that $x < y$. Then there is c with $x < c < y$ so that

$$\frac{f(x) - f(y)}{x - y} = f'(c),$$

and by assumption $f'(c) < 0$. Then (since $x - y$ is now negative), we still get $f(x) > f(y)$, a contradiction. We can treat the minimum case similarly.

Note that under these assumptions, we actually have that x is the unique global maximum (or minimum) on the interval $[a, b]$. Intuitively, this says we get a maximum if f is increasing as we approach x from the left and decreasing as we leave x to the right.

From the First derivative test, we easily obtain:

Second derivative test Let f be a once continuously differentiable function. Let x be a critical point for f. If f is twice differentiable at x and $f''(x) < 0$ then f has a local maximum at x. If f is twice differentiable at x and $f''(x) > 0$ then f has a local minimum at x.

Proof of the second derivative test It suffices to treat the maximum case as the minimum case proceeds similarly. By the definition of the derivative, we have

$$f'(y) = f'(x) + f''(x)(y - x) + o(y - x).$$

Since $f''(x) < 0$, there must be a small interval around x, so that to the left of x, we have f' positive and to the right, it is negative. We apply the first derivative test.

All of this is closely related to the notion of convexity and concavity.

Definition A function $f(x)$ is concave if it lies above all its secants. Precisely f is *concave* if for any a, b, x with $x \in (a, b)$, we have

$$f(x) \geq \frac{b - x}{b - a} f(a) + \frac{x - a}{b - a} f(b).$$

We say f is *strictly concave* if under the same conditions

$$f(x) > \frac{b - x}{b - a} f(a) + \frac{x - a}{b - a} f(b).$$

Similarly f is convex if it lies below all its secants. Precisely f is *convex* if for any a, b, x with $x \in (a, b)$, we have

$$f(x) \leq \frac{b - x}{b - a} f(a) + \frac{x - a}{b - a} f(b).$$

We say f is *strictly convex* if under the same conditions

$$f(x) < \frac{b - x}{b - a} f(a) + \frac{x - a}{b - a} f(b).$$

Theorem Let f be twice continuously differentiable. Then f is concave if and only if for every x, we have $f''(x) \leq 0$, and convex if and only if for every x, we have $f''(x) \geq 0$.

We will leave the proof of the theorem as an exercise but indicate briefly why this is true locally. If $f''(x) \leq 0$, we have

$$f(y) = f(x) + f'(x)(y - x) + \frac{f''(x)}{2}(y - x)^2 + o((y - x)^2).$$

We observe that for a, b close to x, if $f''(x) < 0$, we have that $(a, f(a))$ and $(b, f(b))$ are below the tangent line to the graph of f at x. Thus the point $(x, f(x))$ which is on the tangent line is above the secant between $(a, f(a))$ and $(b, f(b))$.

Concavity has a lot to do with optimization.

Example: Resource allocation problems

Let f and g be two functions which are continuous on $[0,\infty)$ and twice continuously differentiable with strictly negative second derivative on $(0,\infty)$. Let t be a fixed number.

Consider

$$W(x) = f(x) + g(t-x).$$

[Interpretation: You have t units of a resource and must allocate them between two productive uses. The function of W might represent the value of allocating x units to the first use and $t-x$ units to the second use. The concavity represents the fact that each use has diminishing returns.]

Under these assumptions, if $W(x)$ has a critical point in $(0,t)$, then the critical point is a maximum.

If, in addition,

$$\lim_{x \longrightarrow 0} f'(x) = \infty,$$

and

$$\lim_{x \longrightarrow 0} g'(x) = \infty,$$

then we are guaranteed that $W(x)$ has a unique maximum in $[0,t]$. This is because

$$\lim_{x \longrightarrow 0} W'(x) = \infty,$$

and

$$\lim_{x \longrightarrow t} W'(x) = -\infty.$$

By the intermediate value theorem, there is a zero for $W'(x)$ in $(0,t)$. It is unique since $W'(x)$ is strictly decreasing.

Strictly concave functions on $(0,\infty)$ whose derivative converge to ∞ at 0 are ubiquitous in economics. We give an example.

Cobb-Douglas Production function:

The Cobb-Douglas production function gives the output of an economy as a function of its inputs (labor and capital).

$$P(K,L) = cK^{\alpha}L^{1-\alpha}.$$

Here c is a positive constant and α a real number between 0 and 1. The powers of K and L in the function have been chosen so that

$$P(tK,tL) = tP(K,L).$$

That is, if we multiply both the capital and the labor of the economy by t, then we multiply the output by t. Note that if we hold L constant and view $P(K,L)$ as a function of K, then we see this function is defined on $(0,\infty)$ is strictly concave with derivative going to ∞ at 0.

An important principle of economics is that we should pay for capital at the rate of the marginal product of capital. We find this rate by taking the derivative in K and getting $\alpha c K^{\alpha-1} L$. Since we need K units of capital to get the economy to function we pay $\alpha c K^{\alpha} L^{1-\alpha}$. In this way, we see that α represents the share of the economy that is paid to the holders of capital and $1-\alpha$ is the share paid to the providers of labor.

Another way in which optimization can be applied is to prove inequalities.

Arithmetic Geometric mean inequality Let a and b be positive numbers then

$$a^{\frac{1}{2}} b^{\frac{1}{2}} \leq \frac{1}{2}(a+b).$$

This can be proved in a purely algebraic way.

Algebraic proof of arithmetic geometric mean inequality

$$a+b-2a^{\frac{1}{2}} b^{\frac{1}{2}} = (a^{\frac{1}{2}} - b^{\frac{1}{2}})^2.$$

Analytic proof of arithmetic geometric mean inequality

It suffices to prove the inequality when $a+b=1$. This is because

$$(ta)^{\frac{1}{2}} (tb)^{\frac{1}{2}} = t a^{\frac{1}{2}} b^{\frac{1}{2}},$$

while

$$(ta+tb) = t(a+b),$$

so we just pick $t = \frac{1}{a+b}$.

Thus what we need to prove is

$$\sqrt{x}\sqrt{1-x} \leq \frac{1}{2},$$

when $0 < x < 1$. We let

$$f(x) = \sqrt{x}\sqrt{1-x},$$

and calculate

$$f'(x) = \frac{\sqrt{1-x}}{2\sqrt{x}} - \frac{\sqrt{x}}{2\sqrt{1-x}}.$$

$$f''(x) = -\frac{1}{2\sqrt{x}\sqrt{1-x}} - \frac{\sqrt{x}}{4(1-x)^{\frac{3}{2}}} - \frac{\sqrt{1-x}}{4x^{\frac{3}{2}}}.$$

All terms in the last line are negative so f is strictly concave. The unique critical point is at $x = \frac{1}{2}$, where equality holds. We have shown that

$$\sqrt{x}\sqrt{1-x} \leq \frac{1}{2},$$

since $\frac{1}{2}$ is the maximum.

The analytic proof looks a lot messier than the algebraic one, but it is more powerful. For instance, by the same methods, we get that if $\alpha, \beta > 0$ and

$$\alpha + \beta = 1,$$

then

$$a^{\alpha} b^{\beta} \leq \alpha a + \beta b,$$

for $a, b > 0$.