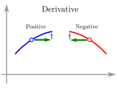
Optimization

< Higher derivatives | Home Page | Differentials >

One of the most important applications of derivatives is *optimization*. In some introductory calculus classes these types of problems are called *max/min problems*: given a function, what is the maximum or minimum output subject to some constraints. This module will review how derivatives can be used in these problems and give some of the reasons why these methods work.

Critical points

First, observe that for a differentiable function f, if the derivative is not zero at a point, then that point cannot be a maximum or a minimum. For instance, if the derivative is positive, then the output is increasing with respect to the input, so by increasing the input, one can increase the output. Hence, the point is not a maximum. If the derivative is negative, then decreasing the input will increase the output, so that point cannot be a maximum.

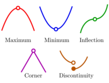


Similarly, a point cannot be a minimum if the derivative is not zero. Thus, the only possible inputs where a maximum or minimum can occur are those where the derivative is zero. This motivates the following definition

Critical point

A *critical point* of a function f is an input $x = a$ where either $f'(a) = 0$ or where the derivative is undefined.

Critical points include maximum and minimum points (called *extrema*) as well as inflection points; these are the points where the derivative is 0. Other critical points occur at corner points or discontinuities, where the derivative is undefined. The reason for including points where the derivative is not defined is that such a point could be a maximum or minimum:



Example

Compute the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$. Hide

The derivative $f'(x) = 3x^2 - 12x + 9$ is defined everywhere, so the critical points are where $f'(x) = 0$. Since

$$\begin{aligned} f'(x) &= 3x^2 - 12x + 9 \\ &= 3(x^2 - 4x + 3) \\ &= 3(x-3)(x-1), \end{aligned}$$

the critical points are $x = 3$ and $x = 1$.

Classifying critical points

Once one has computed a critical points $x = a$, one can classify whether it is a maximum or minimum using the *second derivative test:*

Second Derivative Test

Suppose $x = a$ is a critical point of f where $f'(a) = 0$.

1. If $f''(a) > 0$, then f has a local minimum at a.
2. If $f''(a) < 0$, then f has a local maximum at a.
3. If $f''(a) = 0$, then the test fails.

In the third case, one can use the Taylor expansion about $x = a$ to determine the behavior of the function. In this case, $x = a$ could still be a local maximum, minimum, or inflection point.

Hide

The second derivative test is justified by considering the Taylor series for f about $x = a$:

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots \\ &= f(a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots, \end{aligned}$$

since $f'(a) = 0$. Thus, when x is close to a, $f(x)$ behaves like a parabola centered at $x = a$. Recall that the sign of the coefficient of the square term in a parabola determines if the parabola opens up or down. A positive coefficient means the parabola opens upward, and a negative coefficient means the parabola opens downward.

Here, the coefficient of the $(x-a)^2$ is $\frac{1}{2}f''(a)$. So if $f''(a) > 0$, then the parabola opens upward, meaning $f(a)$ is a local minimum of f. If $f''(a) < 0$, then the parabola opens downward, meaning $f(a)$ is a local maximum of f. If $f''(a) = 0$, then one has to look at more terms of the Taylor series to determine f's behavior at a.

Example

Use the Taylor series about $x = 0$ for

$$\sin^2 x \ln(\cos x)$$

to determine whether the function has a local maximum, local minimum, or inflection point at $x = 0$. (Take as a given that $x = 0$ is a critical point).

Hide

Expanding and multiplying the Taylor series gives

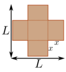
$$\begin{aligned} \sin^2 x \ln(\cos x) &= \left(x + O(x^3)\right)^2 \ln\left(1 - \frac{1}{2!}x^2 + O(x^4)\right) \\ &= (x + O(x^3))(x + O(x^3))\left(-\frac{1}{2}x^2 + O(x^4)\right) \\ &= -\frac{1}{2}x^4 + O(x^6). \end{aligned}$$

Thus, near $x = 0$ the function behaves like $-\frac{1}{2}x^4$, which is downward opening (because of the negative coefficient) and U shaped (because it is an even power). Therefore, the function has a local maximum at $x = 0$.

Note that the second derivative test, besides being tricky to apply with all of the product rules and chain rules, would ultimately be inconclusive in this example.

Example

Consider a square sheet of cardboard of side length L. By cutting equal sized squares of side length x from each corner of the sheet and folding up the flaps which are formed, one gets an open box:



Note that as x gets bigger, the box gets taller but the area of the base of the box shrinks. As x gets smaller, the area of the base grows, but the height shrinks. Find the value of x which maximizes the volume of the resulting box.

Hide

The volume of the box is the area of the base times the height. The base is a square of side length $L - 2x$ (since x has been cut from both sides). The height of the box is x. Thus

$$V = (L-2x)^2 \cdot x = 4x^3 - 4Lx^2 + L^2x$$

Finding the critical points means taking the derivative with respect to x and setting equal to 0:

$$\frac{dV}{dx} = 12x^2 - 8Lx + L^2 = 0.$$

This factors as $(6x - L)(2x - L) = 0$,

so the critical points are $x = \frac{L}{6}$ and $x = \frac{L}{2}$. To apply the second derivative test, we compute

$$\frac{d^2V}{dx^2} = 24x - 8L$$

and evaluate at each critical point:

$$\left.\frac{d^2V}{dx^2}\right|_{x=L/2} = 4L > 0 \quad \text{and} \quad \left.\frac{d^2V}{dx^2}\right|_{x=L/6} = -4L < 0.$$

Thus, $x = \frac{L}{2}$ is a local minimum and $x = \frac{L}{6}$ is a local maximum. (Note also that for $x = \frac{L}{2}$ there is no cardboard left, since the removed corners have consumed the entire square!).

The volume that results from $x = \frac{L}{6}$ is

$$V = \left(\frac{2L}{3}\right)^2 \frac{L}{6} = \frac{2L^3}{27}.$$

Example

Classify the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$. Hide

As found in a previous example, the critical points of f are $x = 3$ and $x = 1$. The second derivative of f is $f''(x) = 6x - 12$. Thus, $f''(3) = 6 > 0$ and $f''(1) = -6 < 0$, and it follows from the second derivative test that 3 is a local minimum of f, and 1 is a local maximum of f.

Example

Suppose a firm producing widgets expects to sell $3000 - 10p^2$ units (where p is the price of the unit). What price p should the firm set to maximize revenue (note that revenue here is just price times quantity sold)? Hide

Revenue is $R(p) = (3000 - 10p^2)p = 3000p - 10p^3$. Taking the derivative gives $R'(p) = 3000 - 30p^2$, and setting equal to 0 gives

$$\begin{aligned} 3000 - 30p^2 &= 0 \\ 3000 &= 30p^2 \\ 100 &= p^2 \\ p &= \pm 10. \end{aligned}$$

So R has critical point $p = 10$ (ignore $p < 0$ since price should be positive). The second derivative is $R''(p) = -60p$. Thus $R''(10) = -600 < 0$, and $p = 10$ is a local maximum.

At this price, the revenue is

$$R(10) = 2000 \cdot 10 = 20000.$$

Global Extrema

While a local maximum or minimum is sometimes useful information, what is usually more important is the global maximum and minimum values of a function on a closed interval $[a, b]$ (or subject to some other constraint such as $x \geq 0$). These are called the *global extrema*, or *absolute extrema*, of a function.

Global extrema on the interval $[a, b]$ either occur at critical points of f or at the endpoints of the interval. So in addition to finding the critical points of f in the interval and checking their values, one must also evaluate f at the endpoints of the interval to find the global extrema.

Example

Find the global extrema of $f(x) = x^3 - 6x^2 + 9x - 5$ on the interval $[2, 4]$. Hide

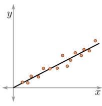
From the prior examples, $x = 1$ and $x = 3$ are the critical points of f. But $x = 1$ is not in the interval $[2, 4]$, so disregard it. Then evaluate f at 2,3,4 to find the extreme values:

$$\begin{aligned} f(2) &= 8 - 24 + 18 - 5 = -3 \\ f(3) &= 27 - 54 + 27 - 5 = -5 \\ f(4) &= 64 - 96 + 36 - 5 = -1. \end{aligned}$$

Thus the absolute maximum is -1 and occurs when $x = 4$. The absolute minimum is -5 and occurs at $x = 3$.

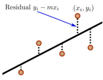
Application: Statistics

In statistics, one often takes experimental data points of the form (x_i, y_i) and looks for a relationship. A very simple relationship is the linear relationship $y = mx$. The data may not follow this relationship perfectly, and there may be some slight experimental error or other noise, so one tries to find the value of m which best fits the data:



This process, called a *linear regression*, can be framed as an optimization problem. But what is the quantity being optimized?

There are several different linear regression models, depending on the quantity being minimized. These different quantities yield different best fit lines. One of the most common models is called *ordinary least squares*. This method seeks to minimize the sum of the squares of the *residuals*, which are the vertical distances from the points to the line:



As shown above, the residual for a given point (x_i, y_i) is $y_i - mx_i$. Thus, the quantity being minimized is

$$S(m) = \sum_i (y_i - mx_i)^2.$$

Taking the derivative with respect to m gives

$$\begin{aligned} \frac{dS}{dm} &= \sum_i 2(y_i - mx_i)(-x_i) \\ &= \sum_i (-2x_iy_i + 2mx_i^2) \\ &= -2\sum_i x_iy_i + 2m\sum_i x_i^2. \end{aligned}$$

Setting this equal to 0 and solving for m gives

$$m = \frac{\sum_i x_iy_i}{\sum_i x_i^2}.$$

Applying the second derivative test, we compute

$$\frac{d^2S}{dm^2} = 2\sum_i x_i^2 > 0,$$

so the above value of m minimizes the sum of squares (hence the *least squares* name).

NOTE

To find the line of best fit of the form $y = mx + b$ requires methods of multivariable calculus (because there are two variables, m and b, which need to be optimized). Optimization with multiple variables is not much more difficult than for a single variable, but these methods are beyond the scope of this course.

EXERCISES

- Find all the local maxima and minima of the function $y = xe^{-x^2}$.
- Which type of critical point does the function $f(x) = e^{\sin(x^4)}\cos(x^2)$ have at zero ?
- Use a Taylor series about $x = 0$ to determine whether the function $f(x) = \sin^3(x^3)$ has a local maximum or local minimum at the origin.
- Find the location of the global maximum and minimum of $f(x) = x^3 - 6x^2 + 1$ on the interval $[-1, 7]$.
- Consider a stretch of highway in which cars are traveling at an average speed v. The "traffic density" u is the total amount of cars on our stretch of road divided by its length. These two quantities are related: the less cars on the road, the faster drivers are able to go. On the other hand, if traffic becomes heavy, drivers will naturally decrease their speed. The so-called "parabolic model" assumes that this relationship is dictated by the equation: $u = u_{\max}\left(1 - \frac{v}{v_{\max}}\right)$ where $u_{\max}$ represents the capacity of the road, and $v_{\max}$ the speed limit on it. The amount of cars passing through our road is called the "traffic flux" or "throughput", and is given by the product of the traffic density and the average speed: $F = uv$. Using the parabolic model, find out at what average speed v_* the flux through our road is maximized.
- A manufacturing company wants to know how many workers it should hire. If it employs too many people, the machines in the factory will be overutilized and the workers will have to wait until they are free, thus reducing the number of units each one will produce in a day's work. On the other hand, too few workers would leave the machines idle for long periods of time. A rough model for the relationship between the number n of workers and their productivity p is given by the equation $p = p_{\max}\left(1 - \frac{n}{n_{\max}}\right)$ where $p_{\max} = 10$ is the maximum number of units a worker can produce in a day and $n_{\max} = 100$ is the maximum number of workers the factory can accommodate. The amount of units U manufactured in the whole factory in one day is equal to the product of the number of workers and the number of units each one produces: $U = np$. How many workers should the company hire in order to maximize its production?
- A technology company has just invented a new gadget. In order to maximize the profit derived from its sale, the company must make a critical decision: at what price should it be sold? A market study suggests that the number N of units sold would approximately follow the equation $N = N_{\max}e^{-P/\lambda}$, where P is the sale price, $N_{\max} = 10,000,000$ is the number of units that would saturate the market, and $\lambda = \$50$. If it costs $\$250$ to manufacture one of these gadgets, at what price P_* would be profit of the company be maximized?
- The manufacturing process of a certain chemical substance is exothermic, that is, it releases heat. The amount of heat released, Q, depends on the temperature T at which the process is carried out, and it is given by the equation $Q = \alpha(T - T_0)^{-2}e^{(T-T_0)/\lambda}$, where $T_0 = 70^oF$ is the room temperature of the manufacturing plant, and $\alpha = 3000\ J\ (^oF)^2$ and $\lambda = 50^oF$. If the temperature T must be maintained above $100^o F$, at what temperature T_* would be the heat loss be minimized?

< Higher derivatives | Home Page | Differentials >