TAYLOR'S POLYNOMIALS IN 1 OR MORE VARIABLES

ORI GUREL-GUREVICH

1. Single Variable

Definition 1.1. The n-th order Taylor's Polynomial of f around x_0 is

$$T_n(x) =$$
$$f(x_0)+f'(x_0)(x-x_0)+f''(x_0)\frac{(x-x_0)^2}{2}+f'''(x_0)\frac{(x-x_0)^3}{6}+\cdots+f^{(n)}\frac{(x-x_0)^n}{n!}.$$

Example 1.2. $T_1(x) = f(x_0) + f'(x_0)(x - x_0)$ is the line tangent to f at $(x_0, f(x_0))$.

Example 1.3. If $f(x) = \log(x)$ (where log is the natural logarithm - base e) then $f'(x) = 1/x$ and $f''(x) = -1/x^2$. The 2nd order Taylor Polynomial around 1 is $T_2(x) = 0 + (x-1) - \frac{1}{2}(x-1)^2$.

This T_n has the property that its value and first n derivatives at the point X_0 coincide with the value of f and its first n derivatives at x_0, causing T_n to behave very similarly to f near x_0. To see this, notice that $g(x) = \frac{(x-x_0)^k}{k!}$ has the following properties:

$$g^{(\ell)}(x) = \begin{cases} \frac{(x-x_0)^{k-\ell}}{(k-\ell)!} & \ell < k \\ 1 & \ell = k \\ 0 & \ell > k \end{cases}$$

In particular,

$$g^{(\ell)}(x_0) = \begin{cases} 1 & \ell = k \\ 0 & \ell \neq k \end{cases}$$

So, the k-th term in the definition of $T_n(x)$ contribute exactly $f^{(k)}$ to the k-th derivative of T_n at x_0 and nothing to the other derivatives.

Plenty of other examples and explanations can be found in the book (chapter 12.10 and 12.11).

Definition 1.4. The *remainder* or *error* of T_n is defined to be $R_n(x) = f(x) - T_n(x)$, so we can write $f(x) = T_n(x) + R_n(x)$.

We would like to know that this error is small, meaning that T_n is a good approximation for f. For this we have Taylor's inequality:

Theorem 1.5 (Taylor's Inequality)**.** *If for any x such that $|x-x_0| \le d$ we have $f^{(n+1)}(x) \le M$ than for any x such that $|x-x_0| \le d$ we have*

$$|R_n(x)| \le M\frac{|x-x_0|^{n+1}}{(n+1)!}.$$

Example 1.6. If $f(x) = \log(x)$ as before, then $f^{(3)} = 2/x^3$. If we restrict ourselves to $|x-1| \le \frac{1}{2}$ than $|f^{(3)}| \le 16$, since $f^{(3)}$ is decreasing so its maximum on $[\frac{1}{2}, \frac{3}{2}]$ is achieved at $\frac{1}{2}$. Hence, the remainder $R_2(x) \le 16(x-1)^3/6$ for any $\frac{1}{2} \le x \le \frac{3}{2}$. This means that we can estimate $\log(1.1)$ by $T_2(1.1) = (1.1-1) - \frac{1}{2}(1.1-1)^2 = 0.1-0.005 = 0.095$ and the error $R_2(1.1) = \log(1.1) - 0.095$ is at most $16(1.1-1)^3/6 = 0.00266...$ In actuality, $\log(1.1) = 0.09531...$

2. Multiple Variables

The 1st order Taylor's Polynomial of f around (x_0, y_0) is

$$T_1(x) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0).$$

This is exactly the function describing the tangent plane at $(x_0, y_0, f(x_0, y_0))$. In other words, $z = T_1(x)$ is the equation of this tangent plane.

The 2nd order Taylor's Polynomial of f around (x_0, y_0) is

$$T_2(x) = T_1(x) + \frac{f_{xx}(x_0,y_0)}{2}(x-x_0)^2 + f_{xy}(x_0,y_0)(x-x_0)(y-y_0) + \frac{f_{yy}(x_0,y_0)}{2}(y-y_0)^2.$$

Recall that for "reasonable" functions Clairaut's theorem apply and we get that $f_{xy} = f_{yx}$ explaining why only one of them appear in the formula (for functions where $f_{xy} \ne f_{yx}$ we will generally not be interested in T_2 since it won't be a good approximation).

In general, The n-th order Taylor's Polynomial of f around (x_0, y_0) is

$$T_n(x,y) = \sum_{k=0}^{n}\sum_{\ell=0}^{n-k} \frac{\partial^k\partial^\ell f}{\partial x^k \partial y^\ell}(x_0,y_0)\frac{(x-x_0)^k(y-y_0)^\ell}{k!\ell!}$$

For more than 2 variables the formula is similar. For example, in 3 variables, the terms would be of the form

$$\frac{\partial^k\partial^\ell\partial^m f}{\partial x^k \partial y^\ell \partial z^m}(x_0,y_0,z_0)\frac{(x-x_0)^k(y-y_0)^\ell(z-z_0)^m}{k!\ell!m!}$$

and we would sum over all indices such that $k+\ell+m \leq n$.

Example 2.1. Let $f(x,y,z)=e^{xyz}$. Then calculating all derivatives up to 3rd order, we find that the only nonzero one at $(0,0,0)$ is $f_{xyz}(0,0,0)=1$. Therefore, $T_3(x,y,z)=1+xyz$.

We will focus mostly on the case 2nd order Taylor Polynomial of a 2 variables function.

3. Local Minimum and Maximum

The next definitions and theorem work for any n and any number of variables. For the 1 variable case interpret P_0 as x_0, for the 2 variables case, interpret P_0 as (x_0,y_0), etc.

Definition 3.1. A point P_0 is a *local maximum* for f, if there is a small disc around P_0, such that the value of f at any point in that disc is at most $f(P_0)$.

Definition 3.2. A point P_0 is a *strict local maximum* for f, if there is a small disc around P_0, such that the value of f at any point in that disc is strictly less than $f(P_0)$.

Definition 3.3. A point P_0 is a *local minimum* for f, if there is a small disc around P_0, such that the value of f at any point in that disc is at least $f(P_0)$.

Definition 3.4. A point P_0 is a *strict local minimum* for f, if there is a small disc around P_0, such that the value of f at any point in that disc is strictly more than $f(P_0)$.

A strict local maximum is a local maximum, and a strict local minimum is a local minimum.

Example 3.5. The function $f(x,y)=x^2+y^2$ has a local minimum at $(0,0)$. It is a strict local minimum.

Example 3.6. The function $g(x,y)=1-x^2$ has a local maximum at $(0,0)$, but it is not a strict local maximum since all the points of the form $(0,y)$ also have $f(0,y)=1=f(0,0)$, and there are points like this in every disc around $(0,0)$ no matter how small the disc is.

The following theorem is what makes Taylor Polynomials useful in the study of local minimum and maximum points.

Theorem 3.7. *Suppose f has all $n+1$-th order derivatives around P_0. If T_n is the n-th order Taylor's Polynomial of f around the point P_0 and P_0 is a strict local maximum for T_n then it is also a strict local maximum for f. If P_0 is **not** a local maximum for T_n then it is also not a local maximum for f. These two implications also hold for minimum instead of maximum.*

Note the difference between the two implications. The first require P_0 to be a strict local maximum, and the second require P_0 to not be a local maximum. The idea behind the theorem is this: for a point P near P_0 the error $T_n(P) - f(P_0)$ is significantly smaller than $|P - P_0|^n$ (the distance between P and P_0 to the n-th power). This means that the error is also small compared to $T_n(P) - f(P_0)$ which is at least of order $|P - P_0|^n$ (this is where the strict max/min requirement comes into play). Hence the sign of $T_n(P) - f(P_0)$ is the same as the sign of $f(P) - f(P_0)$ for P close enough to P_0. In particular, if P_0 is a strict local minimum for T_n than $T_n(P) > T_n(P_0) = f(P_0)$ for all P close enough to P_0 and therefore $f(P) > f(P_0)$ for all these P's, i.e. P_0 is a strict local minimum for f.

Let's consider the implication in 1 variable. If f is a 1 variable function and the first $n-1$ derivatives of f at x_0 are all 0, and $f^{(n)}(x_0) \neq 0$ then the n-th order Taylor's Polynomial of f around x_0 is $T_n(x) = f(x_0) + f^{(n)}(x_0)\frac{(x-x_0)^n}{n!}$. It is straightforward to see that if n is even than T_n has a strict local minimum or maximum, depending on whether $f^{(n)}(x_0)$ is positive or negative, and if n is odd than x_0 is neither a local maximum nor minimum for T_n. Combining with theorem 3.7 we get the following generalization of the second derivative test for functions of 1 variable.

Theorem 3.8 (1-Variable Higher Derivative Test)**.** *Assume $f(x)$ has an $n+1$-th derivative at x_0 and $f^{(k)}(x_0) = 0$ for all $k < n$ and $f^{(n)}(x_0) \neq 0$. If n is even and $f^{(n)}(x_0) > 0$ then x_0 is a strict local minimum for f. If n is even and $f^{(n)}(x_0) < 0$ then x_0 is a strict local maximum for f. If n is odd then x_0 is neither a maximum nor a minimum for f.*

Example 3.9. Let $f(x) = \cos(x) + \frac{x^2}{2}$ and consider $x_0 = 0$. Then the first nonzero derivative is $f^{(4)}(0) = 1$. Hence, 0 is a strict local minimum of f.

Before going to the 2 variables, second order case, we shall give a couple of example in more variables or higher order.

Example 3.10. Suppose that for some $f(x,y,z)$, the 2nd order Taylor Polynomial around $(0,0,0)$ is $T_2(x,y,z) = 7+x^2+4y^2-z^2$. Then $(0,0,0)$ is not a local minimum nor maximum for T_2 and therefore isn't a local minimum nor maximum for f as well. If we had $T_2(x,y,z) = 3-2xy$, then still $(0,0,0)$ is not a local minimum nor maximum for T_2 and so for f. If instead we had $T_2(x,y,z) = 3-2x^2-y^2-4z^2$ then $(0,0,0)$ is a strict local maximum for T_2 and hence for f.

Example 3.11. Suppose that for some $f(x,y)$, the 3rd order Taylor Polynomial around $(1,2)$ is $T_3(x,y) = (x-1)^2+2(x-1)^2(y-2)+4(y-2)^3$, then $(1,2)$ is neither a local maximum nor minimum for T_3 (Why? hint: check what happens when $x=1$) and so it isn't for f.

Finally, some examples about the 2nd order, 2 variables case.

Example 3.12. Let $f(x,y) = e^{-x^2-y^2}$. Then the 2nd order Taylor Polynomial around $(0,0)$ is $T_2(x,y) = 1-x^2-y^2$ which has a strict local maximum at $(0,0)$ and the 3rd order derivatives all exist. Therefore, f also has a strict local maximum at $(0,0)$.

Example 3.13. Let $f(x,y) = \cos(x^2-y)$. Then the 2nd order Taylor Polynomial around $(1,1)$ is

$$
\begin{aligned}
T_2(x,y) &= 1-2(x-1)^2+2(x-1)(y-1)-\frac{1}{2}(y-1)^2 \\
&= 1-2((x-1)-\frac{1}{2}(y-1))^2.
\end{aligned}
$$

This function has a local maximum at $(1,1)$, but it is not strict since all points on the line $(x-1)-\frac{1}{2}(y-1)=0$ give the same value. Hence we cannot conclude from this information alone that f has a maximum there. However, it does have a local maximum since $cos(1^2-1)=1$ is the maximum value cos can attain. Indeed, we can see that $(1,1)$ (and any point on the parabola $x^2-y=0$) is a local maximum for f, but it is not strict.

We see that the question we now face is given a polynomial, find whether a given point is a local maximum or minimum and whether it is strict. We will focus on the 2 variables, second order case. Without loss of generality, we may assume that the point of interest is $(0,0)$ and the polynomial is of the form $T_2(x,y) = Ax^2 + Bxy + Cy^2$, for if we had a linear term ax or by we would immediately know that $(0,0)$ is neither a local maximum nor minimum. Assuming $A \neq 0$ we may rewrite T_2 as

$$T_2(x,y) = A\Big(x^2 + \frac{B}{A}xy + \frac{C}{A}y^2\Big)$$
$$= A\Big(\big(x + \frac{B}{2A}y\big)^2 + \big(\frac{C}{A} - \frac{B^2}{4A^2}\big)y^2\Big).$$

Hence, if $\frac{C}{A} - \frac{B^2}{4A^2} > 0$ (which is equivalent to $4AC - B^2 > 0$ and $A > 0$ then we have a strict local minimum, for in this case $T_2(x,y)$ is always non-negative and the only way we have $T_2(x,y) = 0$ is if both $x + \frac{B}{2A}y = 0$ and $y = 0$, which only happens when $(x,y) = (0,0)$. Similarly, if $4AC - B^2 > 0$ and $A < 0$ then we have a strict local maximum. If $4AC - B^2 < 0$ then we have a saddle point, regardless of the value of A. In this case T_2 has the shape of a hyperbolic paraboloid. Finally, if $4AC - B^2 = 0$ then we either have that T_2 is constant or it describes the shape of a cylinder of a parabola (this cylinder is not necessarily in the direction of one of the axes). If it is a constant, then we learn nothing about f. If it is a cylinder of an upward going parabola ($A > 0$, than $(0,0)$ is a local minimum, but not a strict one (see example 3.13). More importantly, $(0,0)$ is **not** a local maximum for T_2 and hence not for f, so in this case, we did learn something about the behaviour of f near $(0,0)$.

For a Taylor Polynomial we have $A = f_{xx}(0,0)/2$, $B = f_{xy}(0,0)$ and $C = f_{yy}(0,0)/2$, we get the following criteria:

Theorem 3.14 (2-Variables Second Derivative Test)**.** *Assume the second partial derivatives of f exist and are continuous in a small disc around (x_0,y_0) and that $f_x(x_0,y_0) = f_y(x_0,y_0) = 0$. Let*

$$D = f_{xx}(x_0,y_0)f_{yy}(x_0,y_0) - (f_{xy}(x_0,y_0))^2.$$

If $D > 0$ and $f_{xx}(x_0,y_0) > 0$ then (x_0,y_0) is a strict local minimum for f.
If $D > 0$ and $f_{xx}(x_0,y_0) < 0$ then (x_0,y_0) is a strict local maximum for f.
If $D < 0$ then (x_0,y_0) is neither a local minimum nor maximum for f.

Examples in abundance can be found in the textbook.