Hessians and the second derivative test

Sean Fitzpatrick

June 13, 2018

We give here a proof of the second derivative test in general. The proof requires the use of Taylor's theorem for a function of several variables, which we will not prove, and a bit of terminology from linear algebra. The details of the proof are borrowed from the text *Vector Calculus*, 4th edition, by Marsden and Tromba.

Warning: there are a number of instances below where we define a function variables $x_1, \ldots, x_n$ in terms of the vector $\mathbf{x} = \langle x_1, \ldots, x_n \rangle$ via matrix multiplication. In order to get things to come out correctly, we need to think of $\mathbf{x}$ as a **column vector** $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, so that for a vector $x \in \mathbb{R}^n$ and an $m \times n$ matrix A, the function $f(\mathbf{x}) = A\mathbf{x}$ produces a vector in $\mathbb{R}^m$ via the usual matrix multiplication. (The line above probably suggests a fairly convincing typographical reason for using the Stewart-style angle bracket notation, even if this isn't compatible with matrix multiplication.)

1 Taylor polynomials

Before getting into the proof, let's take a brief detour and discuss Taylor polynomials. One way of thinking about differentiability of a function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is to think of the linearization $L(\mathbf{x})$ as the degree one Taylor polynomial

$$P_1(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = f(\mathbf{a}) + \frac{\partial f}{\partial x_1}(\mathbf{a})(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(x_n - a_n).$$

The requirement of differentiability is then that the remainder $R_1(\mathbf{x}) = f(\mathbf{x}) - P_1(\mathbf{x})$ goes to zero faster than $\|\mathbf{x} - \mathbf{a}\|$; that is,

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{R_1(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|} = 0.$$

This says roughly that f and P_1 "agree to first order": we can think of $R_1(\mathbf{x})$ as a function such that $|R_1(\mathbf{x})| \sim C\|\mathbf{x} - \mathbf{a}\|^l$, where C is some constant and $l > 1$. (We are not saying that R_1 is equal to a function of this form, only that its dependence on $\|\mathbf{x} - \mathbf{a}\|$ as $\|\mathbf{x} - \mathbf{a}\| \to 0$ is of an order greater than one.) From here we can go on and ask for degree k Taylor

polynomials $P_k(\mathbf{x})$ that give an "k^{th}-order approximation" of f near $\mathbf{a}$. In other words, we want a polynomial

$$P_k(x_1, \ldots, x_n) = a_0 + a_1 x_1 + \cdots + a_n x_n + a_{11} x_1^2 + a_{12} x_1 x_2 + \cdots + a_{nn} x_n^2$$
$$+ \cdots + a_{1\ldots1} x_1^k + a_{1\ldots12} x_1^{k-1} x_2 + \cdots + a_{n\cdots n} x_n^k$$

in n variables of degree k such that the remainder $R_k(\mathbf{x}) = f(\mathbf{x}) - P_k(\mathbf{x})$ satisfies $R_k(\mathbf{x}) \sim C\|\mathbf{x} - \mathbf{a}\|^l$, with $l > k$. In terms of limits, this means

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{R_k(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|^k} = 0.$$

You've probably already noticed a problem with talking about higher-order polynomials in several variables: the notation gets really messy, since there are so many more possible terms! For example, even a relatively simple case like a degree 3 polynomial in 3 variables looks like

$$P(x, y, z) = a + bx + cy + dz + ex^2 + fxy + gxz + hy^2 + kyz + lz^2$$
$$+ mx^3 + nx^2 y + oxy^2 + pxyz + qx^2 z + rxz^2 + sy^3 + ty^2 z + uyz^2 + vz^3$$

for constants $a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v$! Usually we get around this using "multi-index" notation. We let $\alpha = (a_1, \ldots, a_n)$ denote a n-tuple of non-negative integers, where we define $\mathbf{x}^\alpha = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, $|\alpha| = a_1 + \cdots + a_n$ (so that $\mathbf{x}^\alpha$ is a monomial of order $|\alpha|$), and we denote a possible coefficient of $\mathbf{x}^\alpha$ by a_α. A general k^{th}-order polynomial then looks like

$$P_k(\mathbf{x}) = \sum_{|\alpha|=0}^{k} a_\alpha x^\alpha.$$

For example, in 3 variables, the terms where $|\alpha| = 3$ would involve $\alpha = (3, 0, 0)$, $(2, 1, 0)$, $(2, 0, 1)$, $(1, 2, 0)$, $(1, 0, 2)$, $(0, 3, 0)$, $(0, 2, 1)$, $(0, 1, 2)$, $(0, 0, 3)$, so in the above polynomial $m = a_{(3,0,0)}$, $n = a_{(2,1,0)}$, etc., with $\mathbf{x}^{(3,0,0)} = x^3$, $\mathbf{x}^{(2,1,0)} = x^2 y$, and so on. (Note that $\alpha = (0, \ldots, 0)$ is the only multi-index with $|\alpha| = 0$).

With all of that notational unpleasantness out of the way, we can say what the kth-order Taylor polynomial for f near $\mathbf{a}$ should be: Taylor's Theorem, generalized to n variables, states that

$$P_k(\mathbf{x}) = \sum_{|\alpha|=0}^{k} \frac{f^{(\alpha)}(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha,$$

where $\alpha! = a_1! a_2! \cdots a_n!$, and

$$f^{(\alpha)}(\mathbf{a}) = \left(\frac{\partial^{a_1}}{\partial x_1^{a_1}} \frac{\partial^{a_2}}{\partial x_2^{a_n}} \cdots \frac{\partial^{a_n}}{\partial x_n} f \right)(\mathbf{a}).$$

I'll leave it as an exercise to check that putting $k = 1$ reproduces the linearization $P_1(\mathbf{x})$ (note that if $|\alpha| = 1$ we have to have $\alpha = (1, 0, \ldots, 0)$, $(0, 1, 0, \ldots, 0)$, etc.), and that putting $k = 2$ gives the quadratic approximation discussed below.

2 Quadratic functions

Let A be an $n \times n$ matrix. We say that A is *symmetric* if $A^T = A$, or equivalently, if $a_{ij} = a_{ji}$ for each i, j between 1 and n. To each such A we can associate a *quadratic function* $q : \mathbb{R}^n \to \mathbb{R}$ given by

$$q(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x} = \mathbf{x}^T A\mathbf{x},$$

or in terms of components,

$$q(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x^i x^j.$$

We say that A is *non-degenerate* if $\det A \neq 0$; this is equivalent to saying that A is invertible, or that $A\mathbf{x} = \mathbf{0}$ is possible only if $\mathbf{x} = \mathbf{0}$. (Note however that the corresponding property does not hold for q: it is possible to have $q(\mathbf{x}) = 0$ for $\mathbf{x} \neq \mathbf{0}$ even if the corresponding matrix A is non-degenerate. For example, the quadratic function $q(x, y) = x^2 - y^2$ has $q(1, 1) = 0$ and corresponds to the non-degenerate matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.)

A quadratic function q is called *positive-definite* if $q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$, and $q(\mathbf{x}) = 0$ only for $\mathbf{x} = \mathbf{0}$. (Note that the quadratic function $q(x, y) = x^2 - y^2$ given above is *not* positive definite; however, $\tilde{q}(x, y) = x^2 + y^2$ is.) Similarly, q is *negative-definite* if $q(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$ with $q(\mathbf{x}) = 0$ for $\mathbf{x} = \mathbf{0}$ only.

If $q(\mathbf{x}) = \mathbf{x} \cdot A\mathbf{x}$ is positive(negative)-definite, we refer to the corresponding symmetric matrix A as positive(negative)-definite as well. In general it can be difficult to determine when a given quadratic function (or its corresponding matrix) is positive or negative-definite. In the case of a 2×2 matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ we get

$$q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$
$$= a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(c - \frac{b^2}{a}\right)x_2^2,$$

by completing the square. Since we must have $q(x_1, 0) > 0$ if $x_1 \neq 0$, we get $a > 0$, and since $q(0, x_2) > 0$ for $x_2 \neq 0$, it follows that $ac - b^2 = \det A > 0$. Similarly q is negative-definite if $a < 0$ and $\det A > 0$.

For an $n \times n$ matrix, one test is as follows: consider the sequence of $j \times j$ matrices A_j, for $j = 1, \ldots, n$, given by $A_1 = [a_{11}]$, $A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ldots, A_n = A$. (i.e. we take upper-left square sub-matrices of increasing size.) Then A is positive-definite if and only if $\det A_j > 0$ for each $j = 1, \ldots n$, and A is negative-definite if the signs of $\det A_j$ alternate between negative and positive. (So $\det A_1 = a_{11} < 0, \det A_2 > 0, \det A_3 < 0, \ldots$).

Another approach, which is more illuminating but requires more advanced linear algebra, is to use the fact that for any symmetric matrix A there exists a change of basis such that A becomes a diagonal matrix $\tilde{A}$ with respect to that basis. (i.e. A can be diagonalized.) If the entries $\tilde{a}_{ii}$ of $\tilde{A}$ along the main diagonal are all non-zero, then A is non-degenerate. If they are all positive, then A is positive-definite. If they are all negative, then A is negative-definite.

3

We will need the following lemma below. We can't completely prove this result, since it relies on the Extreme Value Theorem, which we will not prove. (It also requires you to accept that the unit sphere in $\mathbb{R}^n$ is a closed and bounded subset.)

Lemma 2.1. *If $q : \mathbb{R}^n \to \mathbb{R}$ is a positive-definite quadratic function, then there exists a real number $M > 0$ such that*

$$q(\mathbf{x}) \geq M \|\mathbf{x}\|^2$$

for any $x \in \mathbb{R}^n$.

Proof. First, let us consider $q(\mathbf{x})$ on the set B of all $\mathbf{x}$ with $\|\mathbf{x}\| = 1$. The set B is closed and bounded and q is continuous on B, so by the Extreme Value Theorem, q must attain a minimum value M for some $\mathbf{x} \in B$. Now, for any constant $c \in \mathbb{R}$, the fact that q is quadratic implies that $q(c\mathbf{x}) = c^2 q(\mathbf{x})$. For any non-zero $\mathbf{x} \in \mathbb{R}^n$, we know that $\dfrac{\mathbf{x}}{\|\mathbf{x}\|} \in B$, and thus, we have

$$q(\mathbf{x}) = q\left(\|\mathbf{x}\|\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) = \|\mathbf{x}\|^2 q\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) \geq M\|\mathbf{x}\|^2.$$

Finally, if $\mathbf{x} = \mathbf{0}$ we get $q(\mathbf{0}) = 0 = M\|\mathbf{0}\|^2$. $\qquad\qquad\square$

3 The Hessian matrix of a real-valued function

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with continuous second-order partial derivatives. We define the *Hessian matrix* of f at a point $\mathbf{a}$ in the domain of f to be the $n \times n$ symmetric matrix

$$\text{Hess } f(\mathbf{a}) = \frac{1}{2} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{a}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{bmatrix}.$$

Note that Hess $f(\mathbf{a})$ is symmetric by Clairaut's theorem. The factor of $1/2$ is included for convenience with respect to Taylor's theorem. Recall that for a function of one variable, the second-order Taylor polynomial of f about $x = a$ is

$$P_2(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2.$$

For $\mathbf{x} \in \mathbb{R}^n$, let us define the quadratic function $h_{f,\mathbf{a}}(\mathbf{x}) = \mathbf{x} \cdot \text{Hess } f(\mathbf{a})\mathbf{x}$ associated to the Hessian of f at $\mathbf{a}$. Taylor's theorem for functions of several variables tells us that if all the *third* derivatives of f are continuous, then near $\mathbf{a} \in \mathbb{R}^n$ we have

$$f(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a}) + R(\mathbf{a}, \mathbf{x}), \qquad (1)$$

where the remainder term $R(\mathbf{a}, \mathbf{x})$ satisfies

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{R(\mathbf{x}, \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|^2} = 0. \qquad (2)$$

Finally, let us define a critical point $\mathbf{a}$ for f to be *non-degenerate* if Hess $f(\mathbf{a})$ is non-degenerate. Now we're ready to state our result on the second derivative test:

Theorem 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be three times continuously differentiable, and suppose that f has a non-degenerate critical point at $\mathbf{a}$. If $\operatorname{Hess} f(\mathbf{a})$ is positive-definite, then $\mathbf{a}$ is a local minimum for f. Similarly, if $\operatorname{Hess} f(\mathbf{a})$ is negative-definite, then $\mathbf{a}$ is a local maximum for f.*

Proof. Since $\mathbf{a}$ is a critical point for f, we know that $\nabla f(\mathbf{a}) = 0$, and so from (1) we get

$$f(\mathbf{x}) - f(\mathbf{a}) = h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a}) + R(\mathbf{a}, \mathbf{x}).$$

By Lemma 2.1, we have that $h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a}) \geq M \|\mathbf{x} - \mathbf{a}\|^2$, and by (2), there exists a $\delta > 0$ such that for any $\mathbf{x}$ with $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$, we have $|R(\mathbf{a}, \mathbf{x})| < M \|\mathbf{x} - \mathbf{a}\|^2$, by taking $\epsilon = M$ in the definition of the limit. This implies that

$$h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a}) + R(\mathbf{a}, \mathbf{x}) > 0,$$

since $h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a}) \geq M \|\mathbf{x} - \mathbf{a}\|^2 > 0$, (it must be positive since we aren't allowing $\mathbf{x} = \mathbf{a}$) and even if $R(\mathbf{a}, \mathbf{x})$ is negative, its absolute value is less than that of $h_{f,\mathbf{a}}(\mathbf{x} - \mathbf{a})$, so since $\mathbf{x} \neq \mathbf{a}$ the sum of the two terms must be positive. Substituting this into the above equation, we get $f(\mathbf{x}) - f(\mathbf{a}) > 0$ for any $\mathbf{x}$ with $0 < \|\mathbf{x} - \mathbf{a}\| < \delta$, and thus f has a local minimum at $\mathbf{a} \in \mathbb{R}^n$. The case of a local maximum can be handled similarly (or by replacing f with $-f$). $\square$

Notice that we in fact do slightly better than a local minimum: we get the *strict* inequality $f(\mathbf{x}) > f(\mathbf{a})$, and not just $f(\mathbf{x}) \geq f(\mathbf{a})$. Often this fact is expressed by saying that non-degenerate critical points are *isolated* - if f has a critical point at $\mathbf{a}$, then there is some neighbourhood of $\mathbf{a}$ in which f has no other critical points. (We have only established this for local maxima and minima, but this fact is true for non-degenerate critical points in general.) This observation is the starting point for an important area of differential topology known as Morse Theory.