Math 396. The Morse Lemma

1. Motivation

Let V be a finite-dimensional nonzero $\mathbf{R}$-vector space and let $f : U \to \mathbf{R}$ be a C^p-function with $2 \le p \le \infty$. Suppose for $u_0 \in U$ we have $\mathrm{d}f(u_0) = 0$; that is, u_0 is a critical point for f. We seek a convenient coordinate system on a neighborhood of u_0 in U that will help us to see how f behaves near u_0. Just as the second derivative helps us to understand the picture near critical points in the one-variable case (assuming the second derivative doesn't vanish!), namely that the local behavior is concave up or concave down, for the general case we should look at the higher-dimensional second derivative. Recall that for any $u \in U$, the *Hessian* $H_f(u) : V \times V \to \mathbf{R}$ is the bilinear form that is just another name for the total second derivative $D^2 f(u) : V \to \mathrm{Hom}(V, \mathbf{R}) = V^\vee$ (the derivative at $u \in U$ of the C^{p-1} mapping $Df : U \to \mathrm{Hom}(V, \mathbf{R})$ sending $u \in U$ to $Df(u) \in \mathrm{Hom}(V, \mathbf{R})$), and $H_f(u)$ it is a symmetric bilinear form due to "equality of mixed partials".

More concretely, if $\{x_i\}$ is a linear coordinate system dual to a choice of basis of V, then the symmetric bilinear form $H_f(u) : V \times V \to \mathbf{R}$ is described by the symmetric matrix $((\partial_{x_i}\partial_{x_j} f)(u))$ of second-order partials. The associated quadratic form $q_f(u) : V \to \mathbf{R}$ defined by $q_f(u) : v \mapsto H_f(u)(v, v)$ is given in coordinates by

$$[q_f(u)] : (x_1, \dots, x_n) \mapsto \sum_{i,j} (\partial_{x_i}\partial_{x_j} f)(u) x_i x_j;$$

this is the 2nd-order part of the Taylor expansion of f at u (in x_i-coordinates) when $u = 0$. The intrinsic quadratic form $q_f(u)$ on V has a signature (r_u, s_u) with $r_u + s_u \le n = \dim V$, so in suitable linear coordinates *that may depend on* u it can be written as $\sum_{i=1}^{r_u} x_i^2 - \sum_{j=1}^{s_u} x_{r_u+j}^2$. This quadratic form is non-degenerate (i.e., $r_u + s_u = n = \dim V$) if and only if $H_f(u) : V \times V \to \mathbf{R}$ is a perfect pairing, which is to say $\det((\partial_{x_i}\partial_{x_j} f)(u)) \ne 0$. A critical point $u_0 \in U$ for f is *non-degenerate* if $H_f(u_0)$ is non-degenerate. (In treatments that do not give a coordinate-free definition of the Hessian as we have done, one has to carry out the extra step of proving "by hand" that the non-vanishing condition on this determinant is independent of the local coordinates; this is a calculation with the transformation laws for second-order partials under change of coordinates, using the hypothesis that the first-order partials all vanish at the point.)

Non-degeneracy at a critical point u_0 is the generalization of the classical condition of non-vanishing for the second derivative at a critical point in calculus. It is therefore reasonable to expect that in the higher-dimensional case when a critical point is non-degenerate we may be able to describe the local behavior of the function near the critical point. There is a general result, called the *Morse Lemma* (named after M. Morse), that shows how this works. It is a pretty application of the implicit function theorem.

2. Main result

The Morse Lemma in the C^∞ case is this:

Theorem 2.1 (Morse)**.** *Let V be a finite-dimensional vector space and $U \subseteq V$ an open set. Let $f : U \to \mathbf{R}$ be a C^∞ function and suppose f has a non-degenerate critical point at $u_0 \in U$. For a suitable C^∞ coordinate system*

$$\varphi = (x_1, \dots, x_n) : U_0 \to \mathbf{R}^n$$

on an open $U_0 \subseteq U$ around u_0 with $\varphi(u_0) = 0$, the mapping $[f] = f \circ \varphi^{-1} : \varphi(U_0) \to \mathbf{R}$ that is "f in the x_i-coordinates" is given by

$$[f](a_1, \dots, a_n) = \sum_{i=1}^{r} a_i^2 - \sum_{j=1}^{s} a_{r+j}^2$$

with $(r, s) = (r, n - r)$ the signature of the quadratic form $q_f(u_0) : V \to \mathbf{R}$ associated to the symmetric bilinear form $H_f(u_0)$ on V.

Remark 2.2. With better technique, one can weaken the assumption of differentiability on f to be that it is C^p with $p \geq 3$ (rather than C^∞) but the resulting coordinate system (U_0, φ) is merely C^{p-2}.

Remark 2.3. The equality of (r, s) with the signature of $q_f(u_0)$ is automatic, as follows: since $q_f(u_0)$ is a coordinate-independent notion, to compute its signature we may use any C^∞ coordinate system we please. Using the one from φ gives $[f] = \sum_{i=1}^{r} x_i^2 - \sum_{j=1}^{s} x_{r+j}^2$ near the origin, in terms of which we compute $q_{[f]}(0)$ is the quadratic form $\sum_{i=1}^{r} x_i^2 - \sum_{j=1}^{s} x_{r+j}^2$ whose signature is obviously (r, s).

As a special case, when $q_f(u_0)$ is positive-definite (resp. negative-definite) we may use the x_i-coordinate system to see visibly that f has a local minimum (resp. local maximum) at u_0, and that in the indefinite (but still non-degenerate!) case there are specific directions in which the values of f go up from $f(u_0)$ and there are specific directions in which the values of f go down from $f(u_0)$ (i.e., there is a "saddle point" at u_0). Of course, the phenomenon of an indefinite non-degenerate $q_f(u_0)$ cannot happen in the 1-dimensional case, so it is a strictly higher-dimensional occurrence.

Let us first give two corollaries of the Morse Lemma, the first of which is quite striking.

Corollary 2.4. *If $u_0 \in U$ is a non-degenerate critical point of f, then f has no other critical points near u_0.*

Proof. Make a local C^∞ change of coordinates near u_0 via the coordinatization afforded by the Morse Lemma. This reduces us to the trivial verification that for $r + s = n$ the function $\sum_{i=1}^{r} x_i^2 - \sum_{j=1}^{s} x_{r+j}^2$ has only 0 as a critical point. ∎

Rather more special is:

Corollary 2.5. *Keep notation and hypotheses as in the Morse Lemma. Suppose $\dim V = 2$ and u_0 is a critical point of f such that $q_f(u_0)$ is neither positive-definite nor negative-definite, which is to say that it has signature $(1, 1)$. In suitable C^∞ coordinates $\{x', y'\}$ near u_0 we have $[f](a, b) = ab$ for (a, b) near $(0, 0)$.*

Proof. The Morse Lemma gives local C^∞ coordinates in which the C^∞ function becomes $u^2 - v^2$. Pass to the C^∞ coordinate system $\{u + v, u - v\}$. ∎

We shall deduce the Morse lemma from a more general result that is called "separation of variables".

Theorem 2.6. *Let U be an open set in a finite-dimensional $\mathbf{R}$-vector space V, and let $f : U \to \mathbf{R}$ be a C^∞ function. Let $u_0 \in U$ be a non-degenerate critical point for f. There exists a C^∞ coordinate system $\varphi = (x_1, \dots, x_n) : U_0 \to \mathbf{R}^n$ on an open neighborhood of u_0 in U with $\varphi(u_0) = 0$ such that $[f] = f \circ \varphi^{-1}$ is given by $\varepsilon x_1^2 + F$ on $\varphi(U_0) \subseteq \mathbf{R}^n$ with F a C^∞ function of $x_2, \dots, x_n$.*

Remark 2.7. For $n = 1$, this theorem just says that if f is a smooth function near the origin in $\mathbf{R}$ with $f(0) = f'(0) = 0$ but $f''(0) \neq 0$ then $f = \varepsilon k^2$ for $\varepsilon = \pm 1$ and some smooth function k near the origin with $k(0) = 0$ but $k'(0) \neq 0$ (as such an k provides a local C^∞ coordinate near the origin on the real line). Let us prove this special case directly. Since $f(0) = 0$ and f is smooth, $f(t) = tg(t)$ for a smooth function g near the origin. (Recall that we construct g using the Fundamental Theorem of

Calculus: for fixed t we define $h(y) = f(ty)$ for $y \in [0,1]$ and $f(t) = h(1) - h(0) = \int_0^1 h' = tg(t)$ with $g(t) = \int_0^1 f'(ty)\mathrm{d}y$ a smooth function of t by the theorem on differentiation through the integral sign.) Since $g(0) = f'(0) = 0$ so repeating the process gives $f(t) = t^2 G(t)$ with G smooth near the origin. Thus, $G(0) = f''(0) \neq 0$, so if this has the same sign as $\varepsilon = \pm 1$ then $f(t) = \varepsilon t^2 (\varepsilon G)(t)$ with εG a smooth function that is positive at the origin. Hence, it admits a smooth positive square root, so we get the result for f.

The Morse Lemma is an inductive consequence of the preceding theorem. Indeed, working in the x_i-coordinates, since the additive decomposition $\varepsilon x_1^2 + F$ of $[f]$ "separates the variables", the non-degeneracy of the Hessian of $[f]$ at the origin (which is equivalent to that of f at 0) is equivalent to the non-degeneracy of the Hessian of F at the origin in $\mathbf{R}^{n-1}$. But the "separation of variables" also shows that F must be a critical point at u_0 since $[f]$ is, and so induction on n permits us to compose $x_2, \dots, x_n$ with a suitable C^∞ change of coordinates on $\mathbf{R}^{n-1}$ near the origin to make F be a difference of sums of squares of separate coordinates. This gives the desired expression for f in suitable local C^∞ coordinates near u_0.

Remark 2.8. The proof below, if applied to the formulation of separation of variables in the C^p setting, only gives a coordinate change of class C^{p-1}. Hence, if we have finite p then inductively using such a method to try to prove the Morse lemma only gives the result with a coordinate change of class C^{p-n} with $n = \dim V$; in particular, for $p < n$ it gives nothing and the constraint $p \geq n$ forced by our method of proof is very unnatural when p is finite. It is largely for this reason that we restrict attention to the C^∞ case here.

3. Proof of separation of variables

By Remark 2.7, we may assume $n = \dim V > 1$. Additive translation has no effect on derivative maps, nor on Hessians (which are higher derivatives). Thus, we may suppose $u_0 = 0$ in V. Since the symmetric bilinear form $H_f(u_0)$ is nonzero, its associated quadratic form $q_f(u_0) : V \to \mathbf{R}$ is nonzero. By the structure theorem for quadratic spaces over $\mathbf{R}$, we may choose linear coordinates $\{y_1, \dots, y_n\}$ on V such that $q_f(u_0)$ is in standard diagonal form, say $\varepsilon y_1^2 + \dots$ with $\varepsilon = \pm 1$. In particular, $(\partial_{y_1}^2 f)(0) = \varepsilon$. For $|y_i|$ small, consider

$$h(y_1, \dots, y_n) = (\partial_{y_1} f)(y_1, \dots, y_n),$$

so $\partial_{y_1} h = \partial_{y_1}^2 f$ is non-vanishing near the origin (since its value at the origin is $\varepsilon \neq 0$). Since 0 is a critical point for f, clearly $h(0) = 0$. Since also $n > 1$, the implicit function theorem implies that for each $(y_2, \dots, y_n)$ near the origin there exists a unique $g(y_2, \dots, y_n)$ near 0 satisfying

$$h(g(y_2, \dots, y_n), y_2, \dots, y_n) = 0,$$

(so $g(0) = 0$) and g a C^∞ function.

Thus, if we fix $c > 0$ then by continuity of g we conclude that for $|a_2|, \dots, |a_n|$ sufficiently small (depending on c) the function $f(y_1, a_2, \dots, a_n)$ has a unique critical point at $y_1 = g(a_2, \dots, a_n)$ in the interval $(-c, c)$ and the second derivative at this critical point has the same sign as ε. By taking c possibly smaller, we can assume that $|a_2|, \dots, |a_n| < c$ is "sufficiently small". Replacing f with $-f$ if necessary, we may suppose $\varepsilon = 1$, so $f(y_1, a_2, \dots, a_n)$ on $(-c, c)$ has a unique minimum at $y_1 = g(a_2, \dots, a_n)$ with positive second derivative there. Thus, for $a_2, \dots, a_n \in (-c, c)$, the difference

$$y_1 \mapsto f(y_1, a_2, \dots, a_n) - f(g(a_2, \dots, a_n), a_2, \dots, a_n)$$

is non-negative with a unique zero at $y_1 = g(a_2, \dots, a_n)$ and a positive second derivative at this minimum point.

Suppose that we can express the difference

$$k(y_1, \ldots, y_n) = f(y_1, \ldots, y_n) - f(g(y_2, \ldots, y_n), y_2, \ldots, y_n) \geq 0$$

as the square of a C^∞ function h near the origin. By defining the C^∞ function $F(y_2, \ldots, y_n) = f(g(y_2, \ldots, y_n), y_2, \ldots, y_n)$ near the origin we get $f(y_1, \ldots, y_n) = h^2 + F(y_2, \ldots, y_n)$, so we would be done as along as $\{h, y_2, \ldots, y_n\}$ is a C^∞ coordinate system near the origin. By the inverse function theorem, this amounts to the condition that $\partial_{y_1} h$ be nonzero at the origin. But such non-vanishing is clear because for y_1 near 0 we see that

$$h(y_1, 0, \ldots, 0)^2 = f(y_1, 0, \ldots, 0) - f(g(0, \ldots, 0), 0, \ldots, 0) = f(y_1, 0, \ldots, 0)$$

has Taylor expansion $y_1^2 + \ldots$ at the origin (as $f(0) = 0$, $(\partial_{y_1} f)(0) = 0$, and $(\partial_{y_1}^2 f)(0) = \varepsilon = 1$), so the Taylor expansion of $h(y_1, 0, \ldots, 0)$ at the origin must be $\pm y_1 + \ldots$.

It remains to prove that $k(y_1, \ldots, y_n)$ is the square of a C^∞ function near the origin. By the inverse function theorem, $y_1' = y_1 - g(y_2, \ldots, y_n), y_2, \ldots, y_n$ is a C^∞ coordinate system near the origin. If we let K denote k expressed in these coordinates, then $K(y_1', y_2, \ldots, y_n)$ is a C^∞ function near the origin that vanishes for $y_1' = 0$. By applying the fundamental theorem of calculus to $u(t) = K(ty_1', y_2, \ldots, y_n)$ with $y_1', y_2, \ldots, y_n$ all fixed,

$$K(y_1', y_2, \ldots, y_n) = u(1) - u(0) = \int_0^1 (\mathrm{d}u/\mathrm{d}t)\mathrm{d}t = y_1' \int_0^1 (\partial_1 K)(ty_1', y_2, \ldots, y_n)\, \mathrm{d}t$$

with integrand that is C^∞ in $y_1', y_2, \ldots, y_n$ (by differentiation through the integral sign and the C^∞ property of K). Thus, we have made a factorization

$$k(y_1, \ldots, y_n) = (y_1 - g(y_2, \ldots, y_n)) I(y_1, \ldots, y_n) \tag{1}$$

with I a C^∞ function near the origin. Fix $y_2 = a_2, \ldots, y_n = a_n$ with $|a_i| < c$. As we have seen above, $k(y_1, a_2, \ldots, a_n) \geq 0$ has a critical point with *positive* second derivative at its unique minimum $y_1 = g(a_2, \ldots, a_n)$ on $(-c, c)$, with $k(y_1, a_2, \ldots, a_n)$ vanishing at this point, so the Taylor expansion for $k(y_1, a_2, \ldots, a_n)$ at $g(a_2, \ldots, a_n)$ begins in degree 2 with positive coefficient. In particular, by considering Taylor expansions it follows from (1) that $I(y_1, a_2, \ldots, a_n)$ vanishes at $y_1 = g(a_2, \ldots, a_n)$ and has positive derivative at this point. Running through the same integration trick with the fundamental theorem of calculus again, we get

$$I(y_1, y_2, \ldots, y_n) = (y_1 - g(y_2, \ldots, y_n)) J(y_1, \ldots, y_n)$$

with $J(g(y_2, \ldots, y_n), y_2, \ldots, y_n) > 0$ for $y_1, \ldots, y_n$ near the origin. Feeding this into (1) and working with $y_1', y_2, \ldots, y_n$ as the C^∞ coordinates near the origin we have

$$K(y_1', y_2, \ldots, y_n) = {y_1'}^2 \widetilde{J}(y_1', y_2, \ldots, y_n)$$

with $\widetilde{J}(0, \ldots, 0) > 0$. We may therefore extract a C^∞ positive square root of $\widetilde{J}$ near the origin, so indeed K (and thus k) is a square of a C^∞ function near the origin.