

SMR1537/4

# Summer School and Conference on Real Algebraic Geometry and its Applications

(4 - 22 August 2003)

---

## Constrained optimization and quadratic mappings

**Andrei Agrachev**

S.I.S.S.A. - International School for Advanced Studies  
Via Beirut 4  
34014 - Trieste  
Italy

---

These are preliminary lecture notes, intended only for distribution to participants



# Chapter 1

## Second order optimality conditions

### 1.1 Hessian

From the geometric viewpoint the study of conditional optimality can be essentially reduced to the study of boundary of the image of a vector-function. Indeed, let us try to minimize a function  $f_0(u)$  under conditions  $f_i(u) = x_i$ ,  $i = 1, \dots, m$ . Consider a vector-function  $F(u) = (f_0(u), f_1(u), \dots, f_m(u))$ . A point  $\bar{u}$  is a desired minimizer if and only if the intersection of the image of  $F$  with the ray  $\{(f_0(\bar{u}) - s, x_1, \dots, x_m) : s \geq 0\}$  in  $\mathbb{R}^{m+1}$  consists of exactly one point  $(f_0(\bar{u}), x_1, \dots, x_m)$ . In particular, if  $\bar{u}$  is a minimizer, then  $F(\bar{u})$  belongs to the boundary of the image of  $F$ .

We are mainly interested in optimality conditions based on the second differential. Consider the problem in a general setting. Let

$$F : \mathcal{U} \rightarrow M$$

be a smooth mapping, where  $\mathcal{U}$  is an open subset in a Banach space and  $M$  is a smooth  $n$ -dimensional manifold. The first differential

$$D_u F : T_u \mathcal{U} \rightarrow T_{F(u)} M$$

is well defined independently on coordinates. This is not the case for the second differential. Indeed, consider the case where  $u$  is a regular point for  $F$ , i.e., the differential  $D_u F$  is surjective. By implicit function theorem, the mapping  $F$  becomes linear in suitably chosen local coordinates in  $\mathcal{U}$  and  $M$ , thus it has no intrinsic second differential. In the general case, well defined independently of coordinates is only a certain part of the second differential.

The differential of a smooth mapping  $F : \mathcal{U} \rightarrow M$  can be defined via the first order derivative

$$D_u F v = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} F(\varphi(\varepsilon)) \quad (1.1)$$

along a curve  $\varphi : (-\varepsilon_0, \varepsilon_0) \rightarrow \mathcal{U}$  with the initial conditions

$$\varphi(0) = u \in \mathcal{U}, \quad \dot{\varphi}(0) = v \in T_u \mathcal{U}.$$

In local coordinates, this derivative is computed as

$$\frac{dF}{du} \dot{\varphi}, \quad \dot{\varphi} = \dot{\varphi}(0).$$

In other coordinates  $\tilde{q}$  in  $M$ , derivative (1.1) is evaluated as

$$\frac{d\tilde{F}}{du} \dot{\varphi} = \frac{d\tilde{q}}{dq} \frac{dF}{du} \dot{\varphi}.$$

Coordinate representation of the first order derivative (1.1) transforms under changes of coordinates as a tangent vector to  $M$  — it is multiplied by the Jacobian matrix  $\frac{d\tilde{q}}{dq}$ .

The second order derivative

$$\left. \frac{d^2}{d\varepsilon^2} \right|_{\varepsilon=0} F(\varphi(\varepsilon)), \quad (1.2)$$

$$\varphi(0) = u \in \mathcal{U}, \quad \dot{\varphi}(0) = v \in T_u \mathcal{U},$$

is evaluated in coordinates as

$$\frac{d^2 F}{du^2}(\dot{\varphi}, \dot{\varphi}) + \frac{dF}{du} \ddot{\varphi}.$$

Transformation rule for the second order directional derivative under changes of coordinates has the form:

$$\begin{aligned} \frac{d^2 \tilde{F}}{du^2}(\dot{\varphi}, \dot{\varphi}) + \frac{d\tilde{F}}{du} \ddot{\varphi} &= \frac{d\tilde{q}}{dq} \left[ \frac{d^2 F}{du^2}(\dot{\varphi}, \dot{\varphi}) + \frac{dF}{du} \ddot{\varphi} \right] \\ &\quad + \frac{d^2 \tilde{q}}{dq^2} \left( \frac{dF}{du} \dot{\varphi}, \frac{dF}{du} \dot{\varphi} \right). \end{aligned} \quad (1.3)$$

The second order derivative (1.2) transforms as a tangent vector in  $T_{F(u)}M$  only if  $\dot{\varphi} = v \in \text{Ker } D_u F$ , i.e., if term (1.3) vanishes. Moreover, it is determined by  $u$  and  $v$  only modulo the subspace  $\text{Im } D_u F$ , which is spanned by the term  $\frac{dF}{du} \ddot{\varphi}$ .

Thus intrinsically defined is the quadratic mapping

$$\begin{aligned} \text{Ker } D_u F &\rightarrow T_{F(u)}M / \text{Im } D_u F, \\ v &\mapsto \left. \frac{d^2}{d\varepsilon^2} \right|_{\varepsilon=0} F(\varphi(\varepsilon)) \pmod{\text{Im } D_u F}. \end{aligned} \quad (1.4)$$

After this preliminary discussion, we turn to formal definitions.

The *Hessian* of a smooth mapping  $F : \mathcal{U} \rightarrow M$  at a point  $u \in \mathcal{U}$  is a symmetric bilinear mapping

$$\text{Hess}_u F : \text{Ker } D_u F \times \text{Ker } D_u F \rightarrow \text{Coker } D_u F = T_{F(u)}M / \text{Im } D_u F. \quad (1.5)$$

In particular, at a regular point  $\text{Coker } D_u F = 0$ , thus  $\text{Hess}_u F = 0$ . Hessian is defined as follows. Let

$$v, w \in \text{Ker } D_u F$$

and

$$\lambda \in (\text{Im } D_u F)^\perp \subset T_{F(u)}^*M.$$

In order to define the value

$$\lambda \text{Hess}_u F(v, w),$$

take vector fields

$$V, W \in \mathcal{U}, \quad V(u) = v, \quad W(u) = w,$$

and a function

$$a \in C^\infty(M), \quad d_{F(u)}a = \lambda.$$

Then

$$\lambda \text{Hess}_u F(v, w) \stackrel{\text{def}}{=} V \circ W (a \circ F)|_u. \quad (1.6)$$

We show now that the right-hand side does not depend upon the choice of  $V$ ,  $W$ , and  $a$ . The first Lie derivative is

$$W(a \circ F) = \langle d_{F(\cdot)}a, F_*W(\cdot) \rangle,$$

and the second Lie derivative  $V \circ W (a \circ F)|_u$  does not depend on second derivatives of  $a$  since  $F_*W(u) = 0$ . Moreover, the second Lie derivative obviously depends only on the value of  $V$  at  $u$  but not on derivatives of  $V$  at  $u$ . In order to show the same for the field  $W$ , we prove that the right-hand side of the definition of Hessian is symmetric w.r.t.  $V$  and  $W$ :

$$\begin{aligned} (W \circ V(a \circ F) - V \circ W(a \circ F))|_u &= [W, V](a \circ F)|_u = \underbrace{d_{F(u)}a \circ D_u F}_{=\lambda} [W, V](u) \\ &= 0 \end{aligned}$$

since  $\lambda \perp \text{Im } D_u F$ . We showed that the mapping  $\text{Hess}_u F$  given by (1.6) is intrinsically defined independently of coordinates as in (1.5).

**Exercise 1.1.** Show that the quadratic mapping (1.4) defined via the second order directional derivative coincides with  $\text{Hess}_u F(v, v)$ .

If we admit only linear changes of variables in  $\mathcal{U}$ , then we can correctly define the full *second differential*

$$D_u^2 F : \text{Ker } D_u F \times \text{Ker } D_u F \rightarrow T_{F(u)}M$$

in the same way as Hessian (1.6), but the covector is arbitrary:

$$\lambda \in T_{F(u)}^*M,$$

and the vector fields are constant:

$$V \equiv v, \quad W \equiv w.$$

The Hessian is the part of the second differential independent on the choice of linear structure in the preimage.

**Exercise 1.2.** Compute the Hessian of the restriction  $F|_{f^{-1}(0)}$  of a smooth mapping  $F$  to a level set of a smooth function  $f$ . Consider the restriction of a smooth mapping  $F : \mathcal{U} \rightarrow M$  to a smooth hypersurface  $S = f^{-1}(0)$ ,  $f : \mathcal{U} \rightarrow \mathbb{R}$ ,  $df \neq 0$ , and let  $u \in S$  be a regular point of  $F$ . Prove that the Hessian of the restriction is computed as follows:

$$\lambda \text{Hess}_u(F|_S) = \lambda D_u^2 F - d_u^2 f, \quad \lambda \perp \text{Im } D_u F|_S, \quad \lambda \in T_{F(u)}^*M \setminus \{0\},$$

and the covector  $\lambda$  is normalized so that

$$\lambda D_u F = d_u f.$$

## 1.2 Local openness of mappings

A mapping  $F : \mathcal{U} \rightarrow M$  is called *locally open* at a point  $u \in \mathcal{U}$  if

$$F(u) \in \text{int } F(O_u)$$

for any neighborhood  $O_u \subset \mathcal{U}$  of  $u$ . In the opposite case, i.e., when

$$F(u) \in \partial F(O_u)$$

for some neighborhood  $O_u$ , the point  $u$  is called *locally geometrically optimal* for  $F$ .

A point  $u \in \mathcal{U}$  is called *locally finite-dimensionally optimal* for a mapping  $F$  if for any finite-dimensional smooth submanifold  $S \subset \mathcal{U}$ ,  $u \in S$ , the point  $u$  is locally geometrically optimal for the restriction  $F|_S$ .

### 1.2.1 Critical points of corank one

*Corank* of a critical point  $u$  of a smooth mapping  $F$  is by definition equal to corank of the differential  $D_u F$ :

$$\text{corank } D_u F = \text{codim } \text{Im } D_u F.$$

In the sequel we will often consider critical points of corank one. In this case the Lagrange multiplier

$$\lambda \in (\text{Im } D_u F)^\perp, \quad \lambda \neq 0,$$

is defined uniquely up to a nonzero factor, and

$$\lambda \text{Hess}_u F : \text{Ker } D_u F \times \text{Ker } D_u F \rightarrow \mathbb{R}$$

is just a quadratic form (in the case  $\text{corank } D_u F > 1$ , we should consider a family of quadratic forms).

Now we give conditions of local openness of a mapping  $F$  at a corank one critical point  $u$  in terms of the quadratic form  $\lambda \text{Hess}_u F$ .

**Theorem 1.1.** *Let  $F : \mathcal{U} \rightarrow M$  be a continuous mapping having smooth restrictions to finite-dimensional submanifolds of  $\mathcal{U}$ . Let  $u \in \mathcal{U}$  be a corank one critical point of  $F$ , and let  $\lambda \in (\text{Im } D_u F)^\perp$ ,  $\lambda \neq 0$ .*

- (1) *If the quadratic form  $\lambda \text{Hess}_u F$  is sign-indefinite, then  $F$  is locally open at  $u$ .*
- (2) *If the form  $\lambda \text{Hess}_u F$  is negative (or positive), then  $u$  is locally finite-dimensionally optimal for  $F$ .*

*Remark.* A quadratic form is locally open at the origin iff it is sign-indefinite.

*Proof.* The statements of the theorem are local, so we fix local coordinates in  $\mathcal{U}$  and  $M$  centered at  $u$  and  $F(u)$  respectively, and assume that  $\mathcal{U}$  is a Banach space and  $M = \mathbb{R}^n$ .

- (1) Consider the splitting into direct sum in the preimage:

$$T_u \mathcal{U} = E \oplus \text{Ker } D_u F, \quad \dim E = n - 1, \quad (1.7)$$

and the corresponding splitting in the image:

$$T_{F(u)} M = \text{Im } D_u F \oplus V, \quad \dim V = 1. \quad (1.8)$$

The quadratic form  $\lambda \text{Hess}_u F$  is sign-indefinite, i.e., it takes values of both signs on  $\text{Ker } D_u F$ . Thus we can choose vectors

$$v, w \in \text{Ker } D_u F$$

such that

$$\lambda F''_u(v, v) = 0, \quad \lambda F''_u(v, w) \neq 0,$$

we denote by  $F'$ ,  $F''$  derivatives of the vector function  $F$  in local coordinates. Since the first differential is an isomorphism:

$$D_u F = F'_u : E \rightarrow \text{Im } D_u F = \lambda^\perp,$$

there exists a vector  $x_0 \in E$  such that

$$F'_u x_0 = -\frac{1}{2} F''_u(v, v).$$

Introduce the following family of mappings:

$$\begin{aligned}\Phi_\varepsilon &: E \times \mathbb{R} \rightarrow M, & \varepsilon \in \mathbb{R}, \\ \Phi_\varepsilon(x, y) &= F(\varepsilon^2 v + \varepsilon^3 y w + \varepsilon^4 x_0 + \varepsilon^5 x), & x \in E, y \in \mathbb{R},\end{aligned}$$

notice that

$$\text{Im } \Phi_\varepsilon \subset \text{Im } F$$

for small  $\varepsilon$ . Thus it is sufficient to show that  $\Phi_\varepsilon$  is open. The Taylor expansion

$$\Phi_\varepsilon(x, y) = \varepsilon^5 (F'_u x + y F''_u(v, w)) + O(\varepsilon^6), \quad \varepsilon \rightarrow 0,$$

implies that the family  $\frac{1}{\varepsilon^5} \Phi_\varepsilon$  is smooth w.r.t. parameter  $\varepsilon$  at  $\varepsilon = 0$ . For  $\varepsilon = 0$  this family gives a surjective linear mapping. By implicit function theorem, the mappings  $\frac{1}{\varepsilon^5} \Phi_\varepsilon$  are submersions, thus are locally open for small  $\varepsilon > 0$ . Thus the mapping  $F$  is also locally open at  $u$ .

(2) Take any smooth finite-dimensional submanifold  $S \subset \mathcal{U}$ ,  $u \in S$ . Similarly to (1.7), (1.8), consider the splittings in the preimage:

$$S \cong T_u S = L \oplus \text{Ker } D_u F|_S,$$

and in the image:

$$\begin{aligned}M &\cong T_{F(u)} M = \text{Im } D_u F|_S \oplus W, \\ \dim W &= k = \text{corank } D_u F|_S \geq 1.\end{aligned}$$

Since the differential  $D_u F : E \rightarrow \text{Im } D_u F$  is an isomorphism, we can choose, by implicit function theorem, coordinates  $(x, y)$  in  $S$  and coordinates in  $M$  such that the mapping  $F$  takes the form

$$F(x, y) = \begin{pmatrix} x \\ \varphi(x, y) \end{pmatrix}, \quad x \in L, \quad y \in \text{Ker } D_u F|_S.$$

Further, we can choose coordinates  $\varphi = (\varphi_1, \dots, \varphi_k)$  in  $W$  such that

$$\lambda F(x, y) = \varphi_1(x, y).$$

Now we write down hypotheses of the theorem in these coordinates. Since  $\text{Im } D_u F|_S \cap W = \{0\}$ , then

$$D_{(0,0)} \varphi_1 = 0.$$

Further, the hypothesis that the form  $\lambda \text{Hess}_u F$  is negative reads

$$\left. \frac{\partial^2 \varphi_1}{\partial y^2} \right|_{(0,0)} < 0.$$

Then the function

$$\varphi_1(0, y) < 0 \quad \text{for small } y.$$

Thus the mapping  $F|_S$  is not locally open at  $u$ . □



There holds the following statement, which is much stronger than the previous one.

**Theorem 1.2 (Generalized Morse's lemma).** *Suppose that  $u \in \mathcal{U}$  is a corank one critical point of a smooth mapping  $F : \mathcal{U} \rightarrow M$  such that  $\text{Hess}_u F$  is a nondegenerate quadratic form. Then there exist local coordinates in  $\mathcal{U}$  and  $M$  in which  $F$  has only terms of the first and second orders:*

$$\begin{aligned} F(x, v) &= D_u F x + \frac{1}{2} \text{Hess}_u F(v, v), \\ (x, v) \in \mathcal{U} &\cong E \oplus \text{Ker } D_u F. \end{aligned}$$

We do not prove this theorem since it will not be used in the sequel.

### 1.2.2 Critical points of arbitrary corank

The necessary condition of local openness given by item (1) of Theorem 1.1 can be generalized for critical points of arbitrary corank.

Recall that *positive (negative) index* of a quadratic form  $Q$  is the maximal dimension of a positive (negative) subspace of  $Q$ :

$$\begin{aligned} \text{ind}_+ Q &= \max \left\{ \dim L \mid Q|_{L \setminus \{0\}} > 0 \right\}, \\ \text{ind}_- Q &= \max \left\{ \dim L \mid Q|_{L \setminus \{0\}} < 0 \right\}. \end{aligned}$$

**Theorem 1.3.** *Let  $F : \mathcal{U} \rightarrow M$  be a continuous mapping having smooth restrictions to finite-dimensional submanifolds. Let  $u \in \mathcal{U}$  be a critical point of  $F$  of corank  $m$ . If*

$$\text{ind}_- \lambda \text{Hess}_u F \geq m \quad \forall \lambda \perp \text{Im } D_u F, \lambda \neq 0,$$

*then the mapping  $F$  is locally open at the point  $u$ .*

*Proof.* First of all, the statement is local, so we can choose local coordinates and assume that  $\mathcal{U}$  is a Banach space and  $u = 0$ , and  $M = \mathbb{R}^n$  with  $F(0) = 0$ .

Moreover, we can assume that the space  $\mathcal{U}$  is finite-dimensional, now we prove this. For any  $\lambda \perp \text{Im } D_u F$ ,  $\lambda \neq 0$ , there exists a subspace

$$E_\lambda \subset \mathcal{U}, \quad \dim E_\lambda = m,$$

such that

$$\lambda \text{Hess}_u F|_{E_\lambda \setminus \{0\}} < 0.$$

We take  $\lambda$  from the unit sphere

$$S^{m-1} = \left\{ \lambda \in (\text{Im } D_u F)^\perp \mid |\lambda| = 1 \right\}.$$

For any  $\lambda \in S^{m-1}$ , there exists a neighborhood  $O_\lambda \subset S^{m-1}$ ,  $\lambda \in O_\lambda$ , such that  $E_{\lambda'} = E_\lambda$  for any  $\lambda' \in O_\lambda$ , this easily follows from continuity of the form  $\lambda' \text{Hess}_u F$  on the unit sphere in  $E_\lambda$ . Choose a finite covering:

$$S^{m-1} = \bigcup_{i=1}^N O_{\lambda_i}.$$

Then restriction of  $F$  to the finite-dimensional subspace  $\sum_{i=1}^N E_{\lambda_i}$  satisfies the hypothesis of the theorem. Thus we can assume that  $\mathcal{U}$  is finite-dimensional. Then the theorem is a consequence of the following Lemmas 1.1 and 1.2.  $\square$

**Lemma 1.1.** *Let  $F : \mathbb{R}^N \rightarrow \mathbb{R}^n$  be a smooth mapping, and let  $F(0) = 0$ . Assume that the quadratic mapping*

$$Q = \text{Hess}_0 F : \text{Ker } D_0 F \rightarrow \text{Coker } D_0 F$$

*has a regular zero:*

$$\exists v \in \text{Ker } D_0 F \text{ s.t. } Q(v) = 0, \quad D_v Q \text{ surjective.}$$

*Then the mapping  $F$  has regular zeros arbitrarily close to the origin in  $\mathbb{R}^N$ .*

*Proof.* We modify slightly the argument used in the proof of item (1) of Theorem 1.1. Decompose preimage of the first differential:

$$\mathbb{R}^N = E \oplus \text{Ker } D_0 F, \quad \dim E = n - m,$$

then the restriction

$$D_0 F : E \rightarrow \text{Im } D_0 F$$

is one-to-one. The equality  $Q(v) = \text{Hess}_0 F(v) = 0$  means that

$$F_0''(v, v) \in \text{Im } D_0 F.$$

Then there exists  $x_0 \in E$  such that

$$F_0' x_0 = -\frac{1}{2} F_0''(v, v).$$

Define the family of mappings

$$\Phi_\varepsilon(x, y) = F(\varepsilon^2 v + \varepsilon^3 y + \varepsilon^4 x_0 + \varepsilon^5 x), \quad x \in E, \quad y \in \text{Ker } D_0 F.$$

The first four derivatives of  $\Phi_\varepsilon$  vanish at  $\varepsilon = 0$ , and we obtain the Taylor expansion

$$\frac{1}{\varepsilon^5} \Phi_\varepsilon(x, y) = F_0' x + F_0''(v, y) + O(\varepsilon), \quad \varepsilon \rightarrow 0.$$

Then we argue as in the proof of Theorem 1.1. The family  $\frac{1}{\varepsilon^5} \Phi_\varepsilon$  is smooth and linear surjective at  $\varepsilon = 0$ . By implicit function theorem, the mappings  $\frac{1}{\varepsilon^5} \Phi_\varepsilon$  are submersions for small  $\varepsilon > 0$ , thus they have regular zeros in any neighborhood of the origin in  $\mathbb{R}^N$ . Consequently, the mapping  $F$  also has regular zeros arbitrarily close to the origin in  $\mathbb{R}^N$ .  $\square$

**Lemma 1.2.** *Let  $Q : \mathbb{R}^N \rightarrow \mathbb{R}^m$  be a quadratic mapping such that*

$$\text{ind}_- \lambda Q \geq m \quad \forall \lambda \in \mathbb{R}^{m*}, \lambda \neq 0.$$

*Then the mapping  $Q$  has a regular zero.*

*Proof.* We can assume that the quadratic form  $Q$  has no kernel:

$$Q(v, \cdot) \neq 0 \quad \forall v \neq 0. \quad (1.9)$$

If this is not the case, we factorize by kernel of  $Q$ . Since  $D_v Q = 2Q(v, \cdot)$ , condition (1.9) means that  $D_v Q \neq 0$  for  $v \neq 0$ .

Now we prove the lemma by induction on  $m$ .

In the case  $m = 1$  the statement is obvious: a sign-indefinite quadratic form has a regular zero.

Induction step: we prove the statement of the lemma for any  $m > 1$  under the assumption that it is proved for all values less than  $m$ .

(1) Suppose first that  $Q^{-1}(0) \neq \emptyset$ . Take any  $v \neq 0$  such that  $Q(v) = 0$ . If  $v$  is a regular point of  $Q$ , then the statement of this lemma follows. Thus we assume that  $v$  is a critical point of  $Q$ . Since  $D_v Q \neq 0$ , then

$$\text{rank } D_v Q = k, \quad 0 < k < m.$$

Consider Hessian of the mapping  $Q$ :

$$\text{Hess}_v Q : \text{Ker } D_v Q \rightarrow \mathbb{R}^{m-k}.$$

The second differential of a quadratic mapping is the doubled mapping itself, thus

$$\lambda \text{Hess}_v Q = 2 \lambda Q|_{\text{Ker } D_v Q}.$$

Further, since  $\text{ind}_- \lambda Q \geq m$  and  $\text{codim } \text{Ker } D_v Q = k$ , then

$$\text{ind}_- \lambda \text{Hess}_v Q = \text{ind}_- \lambda Q|_{\text{Ker } D_v Q} \geq m - k.$$

By the induction assumption, the quadratic mapping  $\text{Hess}_v Q$  has a regular zero. Then Lemma 1.1 applied to the mapping  $Q$  yields that  $Q$  has a regular zero as well. The statement of this lemma in case (1) follows.

(2) Consider now the second case:  $Q^{-1}(0) = \{0\}$ .

(2.a) It is obvious that  $\text{Im } Q$  is a closed cone.

(2.b) Moreover, we can assume that  $\text{Im } Q \setminus \{0\}$  is open. Indeed, suppose that there exists

$$x = Q(v) \in \partial \text{Im } Q, \quad x \neq 0.$$

Then  $v$  is a critical point of  $Q$ , and in the same way as in case (1) the induction assumption for  $\text{Hess}_v Q$  yields that  $\text{Hess}_v Q$  has a regular zero. By Lemma 1.1,  $Q$  is locally open at  $v$  and  $Q(v) \in \text{int } \text{Im } Q$ . Thus we assume in the sequel that  $\text{Im } Q \setminus \{0\}$  is open. Combined with item (a), this means that  $Q$  is surjective.

(2.c) We show now that this property leads to a contradiction which proves the lemma.

The smooth mapping

$$\frac{Q}{|Q|} : S^{N-1} \rightarrow S^{m-1}, \quad v \mapsto \frac{Q(v)}{|Q(v)|}, \quad v \in S^{N-1},$$

is surjective. By Sard's theorem, it has a regular value. Let  $x \in S^{m-1}$  be a regular value of the mapping  $Q/|Q|$ .

Now we proceed as follows. We find the minimal  $a > 0$  such that

$$Q(v) = ax, \quad v \in S^{N-1},$$

and apply optimality conditions at the solution  $v_0$  to show that  $\text{ind}_- \lambda Q \leq m-1$ , a contradiction.

So consider the following finite-dimensional optimization problem with constraints:

$$a \rightarrow \min, \quad Q(v) = ax, \quad a > 0, \quad v \in S^{N-1}. \quad (1.10)$$

This problem obviously has a solution, let a pair  $(v_0, a_0)$  realize minimum. We write down first- and second-order optimality conditions for problem (1.10). There exist Lagrange multipliers

$$(\nu, \lambda) \neq 0, \quad \nu \in \mathbb{R}, \quad \lambda \in T_{a_0 x}^* \mathbb{R}^m,$$

such that the Lagrange function

$$\mathcal{L}(\nu, \lambda, a, v) = \nu a + \lambda(Q(v) - ax)$$

satisfies the stationarity conditions:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial a} &= \nu - \lambda x = 0, \\ \frac{\partial \mathcal{L}}{\partial v} \Big|_{(v_0, a_0)} &= \lambda D_{v_0} Q|_{S^{N-1}} = 0. \end{aligned} \quad (1.11)$$

Since  $v_0$  is a regular point of the mapping  $Q/|Q|$ , then  $\nu \neq 0$ , thus we can set

$$\nu = 1.$$

Then second-order necessary optimality condition for problem (1.10) reads

$$\lambda \text{Hess}_{v_0} Q|_{S^{N-1}} \geq 0. \quad (1.12)$$

Recall that Hessian of restriction of a mapping is not equal to restriction of Hessian of this mapping, see Exercise 1.2 above.

**Exercise 1.3.** Prove that

$$\begin{aligned} \lambda (\text{Hess}_v Q|_{S^{N-1}})(u) &= 2(\lambda Q(u) - |u|^2 \lambda Q(v)), \\ v &\in S^{N-1}, \quad u \in \text{Ker } D_v Q|_{S^{N-1}}. \end{aligned}$$

That is, inequality (1.12) yields

$$\lambda Q(u) - |u|^2 \lambda Q(v_0) \geq 0, \quad u \in \text{Ker } D_{v_0} Q|_{S^{N-1}},$$

thus

$$\lambda Q(u) \geq |u|^2 \lambda Q(v_0) = |u|^2 a_0 \lambda x = |u|^2 a_0 \nu = |u|^2 a_0 > 0,$$

i.e.,

$$\lambda Q(u) \geq 0, \quad u \in \text{Ker } D_{v_0} Q|_{S^{N-1}}.$$

Moreover, since  $v_0 \notin T_{v_0} S^{N-1}$ , then

$$\lambda Q|_L \geq 0, \quad L = \text{Ker } D_{v_0} Q|_{S^{N-1}} \oplus \mathbb{R}v_0.$$

Now we compute dimension of the nonnegative subspace  $L$  of the quadratic form  $\lambda Q$ . Since  $v_0$  is a regular value of  $\frac{Q}{|Q|}$ , then

$$\dim \text{Im } D_{v_0} \frac{Q}{|Q|} = m - 1.$$

Thus  $\text{Im } D_{v_0} Q|_{S^{N-1}}$  can have dimension  $m$  or  $m - 1$ . But  $v_0$  is a critical point of  $Q|_{S^{N-1}}$ , thus

$$\dim \text{Im } D_{v_0} Q|_{S^{N-1}} = m - 1$$

and

$$\dim \text{Ker } D_{v_0} Q|_{S^{N-1}} = N - 1 - (m - 1) = N - m.$$

Consequently,  $\dim L = N - m + 1$ , thus  $\text{ind}_- \lambda Q \leq m - 1$ , which contradicts the hypothesis of this lemma.

So case (c) is impossible, and the induction step in this lemma is proved.  $\square$

Theorem 1.3 is completely proved.