united nations educational, scientific and cultural organization

international atomic energy agency

the **abdus salam** international centre for theoretical physics

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 we try to minimize a function $f_0(u)$ under conditions $f_i(u) = x_i,\ i = 1, \ldots, m$. Consider a vector-function $F(u) = (f_0(u), f_1(u), \ldots, f_m(u))$. A point $\bar{u}$ is a desired minimizer if and only if the intersection of the image of F with the ray $\{(f_0(\bar{u}) - s, x_1, \ldots, x_m) : s \geq 0\}$ in $\mathbb{R}^{m+1}$ consists of exactly one point $(f_0(\bar{u}), x_1, \ldots, x_m)$. 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

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

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

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

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 $F : \mathcal{U} \to M$ can be defined via the first order derivative

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

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

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

In local coordinates, this derivative is computed as

$$\frac{d\,F}{d\,u}\dot{\varphi}, \qquad \dot{\varphi} = \dot{\varphi}(0).$$

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

$$\frac{d\,\widetilde{F}}{d\,u}\dot{\varphi} = \frac{d\,\tilde{q}}{d\,q}\frac{d\,F}{d\,u}\dot{\varphi}.$$

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 $\dfrac{d\,\tilde{q}}{d\,q}$.

The second order derivative

$$\left.\frac{d^2}{d\,\varepsilon^2}\right|_{\varepsilon=0} F(\varphi(\varepsilon)), \tag{1.2}$$
$$\varphi(0) = u \in \mathcal{U}, \qquad \dot{\varphi}(0) = v \in T_u\mathcal{U},$$

is evaluated in coordinates as

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

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

$$\frac{d^2\,\widetilde{F}}{d\,u^2}(\dot{\varphi}, \dot{\varphi}) + \frac{d\,\widetilde{F}}{d\,u}\ddot{\varphi} = \frac{d\,\tilde{q}}{d\,q}\left[\frac{d^2\,F}{d\,u^2}(\dot{\varphi}, \dot{\varphi}) + \frac{d\,F}{d\,u}\ddot{\varphi}\right] + \frac{d^2\,\tilde{q}}{d\,q^2}\left(\frac{d\,F}{d\,u}\dot{\varphi}, \frac{d\,F}{d\,u}\dot{\varphi}\right). \tag{1.3}$$

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

Thus intrinsically defined is the quadratic mapping

$$\operatorname{Ker} D_uF \to T_{F(u)}M/\operatorname{Im} D_uF,$$
$$v \mapsto \left.\frac{d^2}{d\,\varepsilon^2}\right|_{\varepsilon=0} F(\varphi(\varepsilon)) \mod \operatorname{Im} D_uF. \tag{1.4}$$

After this preliminary discussion, we turn to formal definitions.

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

$$\operatorname{Hess}_u F : \operatorname{Ker} D_u F \times \operatorname{Ker} D_u F \to \operatorname{Coker} D_u F = T_{F(u)}M / \operatorname{Im} D_u F. \tag{1.5}$$

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

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

and

$$\lambda \in (\operatorname{Im} D_u F)^{\perp} \subset T^*_{F(u)}M.$$

In order to define the value

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

take vector fields

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

and a function

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

Then

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

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:

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

since $\lambda \perp \operatorname{Im} D_u F$. We showed that the mapping $\operatorname{Hess}_u F$ 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 $\operatorname{Hess}_u F(v, v)$.

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

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

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:

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

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 $F : \mathcal{U} \to M$ to a smooth hypersurface $S = f^{-1}(0)$, $f : \mathcal{U} \to \mathbb{R}$, $df \neq 0$, and let $u \in S$ be a regular point of F. Prove that the Hessian of the restriction is computed as follows:

$$\lambda \operatorname{Hess}_u (F|_S) = \lambda D^2_u F - d^2_u f, \qquad \lambda \perp \operatorname{Im} D_u F|_S, \quad \lambda \in T^*_{F(u)}M \setminus \{0\},$$

and the covector λ is normalized so that

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

1.2 Local openness of mappings

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

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

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 $D_u F$:

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

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

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

is defined uniquely up to a nonzero factor, and

$$\lambda \operatorname{Hess}_u F : \operatorname{Ker} D_u F \times \operatorname{Ker} D_u F \to \mathbb{R}$$

is just a quadratic form (in the case $\operatorname{corank} D_u F > 1$, 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 $\lambda \operatorname{Hess}_u F$.

Theorem 1.1. *Let $F : \mathcal{U} \to M$ 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 $\lambda \in (\operatorname{Im} D_u F)^{\perp}$, $\lambda \neq 0$.*

(1) *If the quadratic form $\lambda \operatorname{Hess}_u F$ is sign-indefinite, then F is locally open at u.*

(2) *If the form $\lambda \operatorname{Hess}_u F$ 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:

$$T_u \mathcal{U} = E \oplus \operatorname{Ker} D_u F, \qquad \dim E = n - 1, \tag{1.7}$$

and the corresponding splitting in the image:

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

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

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

such that

$$\lambda F_u''(v, v) = 0, \qquad \lambda F_u''(v, w) \neq 0,$$

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

$$D_u F = F_u' : E \to \operatorname{Im} D_u F = \lambda^{\perp},$$

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:

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

notice that

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

for small ε. Thus it is sufficient to show that Φ_ε is open. The Taylor expansion

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

implies that the family $\frac{1}{\varepsilon^5}\Phi_\varepsilon$ is smooth w.r.t. parameter ε 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:

$$S \cong T_u S = L \oplus \operatorname{Ker} D_u \left. F \right|_S ,$$

and in the image:

$$M \cong T_{F(u)} M = \operatorname{Im} D_u \left. F \right|_S \oplus W,$$
$$\dim W = k = \operatorname{corank} D_u \left. F \right|_S \geq 1.$$

Since the differential $D_u F : E \to \operatorname{Im} D_u F$ 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

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

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 $\operatorname{Im} D_u \left. F \right|_S \cap W = \{0\}$, then

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

Further, the hypothesis that the form $\lambda \operatorname{Hess}_u F$ 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 $\left. F \right|_S$ 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 $F : \mathcal{U} \to M$ such that $\operatorname{Hess}_u F$ 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:*

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

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:

$$\operatorname{ind}_+ Q = \max \left\{ \dim L \mid Q|_{L \setminus \{0\}} > 0 \right\},$$
$$\operatorname{ind}_- Q = \max \left\{ \dim L \mid Q|_{L \setminus \{0\}} < 0 \right\}.$$

Theorem 1.3. *Let $F : \mathcal{U} \to M$ 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*

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

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 $\lambda \perp \operatorname{Im} D_u F$, $\lambda \neq 0$, there exists a subspace

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

such that

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

We take λ from the unit sphere

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

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 $\lambda' \operatorname{Hess}_u F$ on the unit sphere in E_λ. 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. □

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

$$Q = \operatorname{Hess}_0 F \,:\, \operatorname{Ker} D_0F \to \operatorname{Coker} D_0F$$

has a regular zero:

$$\exists\, v \in \operatorname{Ker} D_0F \;\; s.t. \;\; Q(v) = 0, \quad D_vQ \;\; surjective.$$

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:

$$\mathbb{R}^N = E \oplus \operatorname{Ker} D_0F, \qquad \dim E = n - m,$$

then the restriction

$$D_0F \,:\, E \to \operatorname{Im} D_0F$$

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

$$F_0''(v, v) \in \operatorname{Im} D_0F.$$

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

$$\Phi_\varepsilon(x, y) = F(\varepsilon^2 v + \varepsilon^3 y + \varepsilon^4 x_0 + \varepsilon^5 x), \qquad x \in E, \quad y \in \operatorname{Ker} D_0F.$$

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

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

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 $\mathbb{R}^N$. □

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

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

Then the mapping Q has a regular zero.

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

$$Q(v, \cdot) \neq 0 \qquad \forall\ v \neq 0. \tag{1.9}$$

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 $Q^{-1}(0) \neq 0$. 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

$$\operatorname{rank} D_v Q = k, \qquad 0 < k < m.$$

Consider Hessian of the mapping Q:

$$\operatorname{Hess}_v Q \,:\, \operatorname{Ker} D_v Q \to \mathbb{R}^{m-k}.$$

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

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

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

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

By the induction assumption, the quadratic mapping $\operatorname{Hess}_v Q$ 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 $\operatorname{Im} Q$ is a closed cone.

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

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

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

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

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

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

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

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

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

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

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

such that the Lagrange function

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

satisfies the stationarity conditions:

$$\frac{\partial \mathcal{L}}{\partial a} = \nu - \lambda x = 0, \tag{1.11}$$

$$\left.\frac{\partial \mathcal{L}}{\partial v}\right|_{(v_0,a_0)} = \lambda D_{v_0} Q|_{S^{N-1}} = 0.$$

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

$$\lambda \operatorname{Hess}_{v_0} Q|_{S^{N-1}} \ge 0. \tag{1.12}$$

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

$$\lambda \left(\operatorname{Hess}_v Q|_{S^{N-1}}\right)(u) = 2(\lambda Q(u) - |u|^2 \lambda Q(v)),$$
$$v \in S^{N-1}, \quad u \in \operatorname{Ker} D_v Q|_{S^{N-1}}.$$

That is, inequality (1.12) yields

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

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.,

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

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

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

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

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

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

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

and

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

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

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

Theorem 1.3 is completely proved.