Solutions to Exercises

D. Liberzon, *Calculus of Variations and Optimal Control Theory*

See the last page for the list of all exercises along with page numbers where they appear in the book.

Chapter 1

1.1

The answer is *no*.

Counterexample: on the (x_1, x_2)-plane, consider the function $f(x) = x_1(1 + x_1) + x_2(1 + x_2)$. Let D be the union of the closed first quadrant $\{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}$ and some curve (e.g, a circular arc) directed from the origin into the third quadrant. The origin $x^* = (0, 0)$ is clearly not a local minimum, because $f(x^*) = 0$ but f is negative for small negative values of x_1 and x_2. However, it is easy to check that the listed conditions are satisfied because the feasible directions are $\{(d_1, d_2) : d_1 \geq 0, d_2 \geq 0\}$ and we have $\nabla f(x^*) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\nabla^2 f(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.

1.2

Example: on the (x_1, x_2)-plane, let $h_1(x) = x_1^2 - x_2$ and $h_2(x) = x_2$. Then D consists of the unique point $x^* = (0, 0)$ which is automatically a minimum of *any* function f over D. The gradients are $\nabla h_1(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and $\nabla h_2(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and they are linearly dependent, hence x^* is not a regular point. It remains to choose any function f whose gradient at x^* is not proportional to $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$—e.g., $f(x) = x_1 + x_2$ works.

See also Example 3.1.1 on pp. 279–280 in [Ber99].

Another example, a little more complicated but also more interesting, is to consider, on the (x_1, x_2)-plane, the functions $h_1(x) = x_2$ and $h_2(x) = x_2 - g(x_1)$ where

$$g(x_1) = \begin{cases} x_1^2 & \text{if } x_1 > 0 \\ 0 & \text{if } x_1 \le 0 \end{cases}$$

Then $D = \{x : x_1 \le 0, x_2 = 0\}$. The point $x^* = (0, 0)$ is not a regular point, and we can again easily choose f for which the necessary condition fails. The interesting thing about this example is that the tangent space to D at x^* is not even a vector space: it is a ray pointing to the left.

1.3

Let's do it for 2 constraints, then it will be obvious how to handle an arbitrary number of constraints. For $d_1, d_2, d_3 \in \mathbb{R}^n$, consider the following map from $\mathbb{R}^3$ to itself:

$$F : \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \mapsto \begin{pmatrix} f(x^* + \alpha_1 d_1 + \alpha_2 d_2 + \alpha_3 d_3) \\ h_1(x^* + \alpha_1 d_1 + \alpha_2 d_2 + \alpha_3 d_3) \\ h_2(x^* + \alpha_1 d_1 + \alpha_2 d_2 + \alpha_3 d_3) \end{pmatrix}.$$

The Jacobian of F at $(0, 0, 0)$ is

$$\begin{pmatrix} \nabla f(x^*) \cdot d_1 & \nabla f(x^*) \cdot d_2 & \nabla f(x^*) \cdot d_3 \\ \nabla h_1(x^*) \cdot d_1 & \nabla h_1(x^*) \cdot d_2 & \nabla h_1(x^*) \cdot d_3 \\ \nabla h_2(x^*) \cdot d_1 & \nabla h_2(x^*) \cdot d_2 & \nabla h_2(x^*) \cdot d_3 \end{pmatrix}.$$

Arguing exactly as in the notes, we know that this Jacobian must be singular for any choice of d_1, d_2, d_3. Since x^* is a regular point and so $\nabla h_1(x^*)$ and $\nabla h_2(x^*)$ are linearly independent, we can choose d_1 and d_2 such that the lower left 2×2 submatrix

$$\begin{pmatrix} \nabla h_1(x^*) \cdot d_1 & \nabla h_1(x^*) \cdot d_2 \\ \nabla h_2(x^*) \cdot d_1 & \nabla h_2(x^*) \cdot d_2 \end{pmatrix}$$

is nonsingular (for example, using the Gram-Schmidt orthogonalization process: choose d_1 aligned with $\nabla h_1(x^*)$ and d_2 in the plane spanned by $\nabla h_1(x^*)$ and $\nabla h_2(x^*)$ to be orthogonal to d_1). Since the Jacobian is singular, its top row must be a linear combination of the bottom two, linearly independent by construction, rows:

$$\nabla f(x^*) \cdot d_i = \lambda_1^* \nabla h_1(x^*) \cdot d_i + \lambda_2^* \nabla h_2(x^*) \cdot d_i, \qquad i = 1, 2, 3.$$

Note that the coefficients λ_1^* and λ_2^* are uniquely determined by our choice of d_1 and d_2, and do not depend on the choice of d_3. In other words, we have

$$\nabla f(x^*) \cdot d_3 = \lambda_1^* \nabla h_1(x^*) \cdot d_3 + \lambda_2^* \nabla h_2(x^*) \cdot d_3 \qquad \forall\, d_3 \in \mathbb{R}^3$$

from which it follows that $\nabla f(x^*) = \lambda_1^* \nabla h_1(x^*) + \lambda_2^* \nabla h_2(x^*)$.

1.4

This is Problem 3.1.3 in [Ber99], page 292 (an easier version appears earlier as Problem 1.1.8, page 19). The function being minimized is $f(x) = |x-y| + |x-z|$. Writing $|x-y|$ as $((x-y)^T(x-y))^{1/2}$, and similarly for $|x-z|$, it is easy to compute that

$$\nabla f(x^*) = \frac{x^*-y}{|x^*-y|} + \frac{x^*-z}{|x^*-z|}.$$

By the first-order necessary condition for constrained optimality, this vector must be aligned with the normal vector $\nabla h(x^*)$. Geometrically, the fact that the two unit vectors appearing in the above formula sum up to a constant multiple of $\nabla h(x^*)$ means that the angles they make with it are equal.

1.5, 1.6

These follow easily from the definitions of the first and second variation by writing down the Taylor expansion of $g(y(x) + \alpha\eta(x))$ around $\alpha = 0$ inside the integral:

$$J(y + \alpha\eta) = \int_0^1 g(y(x) + \alpha\eta(x))dx = \int_0^1 \big(g(y(x)) + g'(y(x))\alpha\eta(x) + \frac{1}{2}g''(y(x))\alpha^2\eta^2(x) + o(\alpha)\big)dx.$$

The second variation is

$$\delta^2 J\big|_y(\eta) = \frac{1}{2}\int_0^1 g''(y(x))\eta^2(x)dx.$$

This example also appears in Section 5.5 of [AF66].

1.7

Let $V = C^0([0,1],\mathbb{R})$ with the 0-norm $\|\cdot\|_0$, let $A = \{y \in V : y(0) = y(1) = 0, \|y\|_0 \le 1\}$, and let $J(y) = \int_0^1 y(x)dx$. It is easy to see that A is bounded, that J is continuous, and that J does not have a global minimum over A because the infimum value of J over A is -1 but it's not achieved for any continuous curve. What's not obvious is that A is closed, because to show this we must show that if a sequence of continuous functions $\{y_k\}$ converges to some function y in 0-norm then the limit y is also continuous. The proof of this goes as follows. To show continuity of y, we must show that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that when $|x_1 - x_2| < \delta$ we have $|y(x_1) - y(x_2)| < \varepsilon$. Let k be large enough so that $\|y_k - y\|_0 \le \varepsilon/3$, and let δ be small enough so that $|y_k(x_1) - y_k(x_2)| < \varepsilon/3$ whenever $|x_1 - x_2| < \delta$ (using continuity of y_k). This gives

$$|y(x_1) - y(x_2)| \le |y(x_1) - y_k(x_1)| + |y_k(x_1) - y_k(x_2)| + |y_k(x_2) - y(x_2)| < \varepsilon$$

and we are done. See also [Rud76, p. 150, Theorem 7.12] or [AF66, p. 103, Theorem 3-11] or [Kha02, p. 655] or [Sut75, p. 120, Theorem 8.4.1].