

# 1 Introduction

- 1.1. Suppose that  $f$  is a  $C^2$  function and  $x^*$  is a point of its domain at which we have  $\nabla f(x^*) \cdot d \geq 0$  and  $d^T \nabla^2 f(x^*) d > 0$  for every nonzero feasible direction  $d$ . Is  $x^*$  necessarily a local minimum of  $f$ ? Prove or give a counterexample.

Recall that a local minimum  $x^*$  of  $f$  is defined as a point in the domain  $D$  where there exists an  $\varepsilon > 0$  such that for all  $x \in D$  satisfying  $|x - x^*| < \varepsilon$  we have

$$f(x^*) \leq f(x).$$

Let us take the function given in the problem statement and expand it in a Taylor series about  $x^*$ ,

$$f(x) = f(x^* + \alpha d) = f(x^*) + \alpha \nabla f(x^*) \cdot d + \frac{1}{2} \alpha^2 d^T \nabla^2 f(x^*) d + o(\alpha^2).$$

Here we have chosen  $|d| = 1$ , and so, with  $x = x^* + \alpha d$ , we have

$$|x - x^*| = |x^* + \alpha d - x^*| = |\alpha d| = \alpha.$$

Choose an  $\alpha$  such that  $|o(\alpha^2)| \leq \frac{1}{2} \alpha^2 d^T \nabla^2 f(x^*) d$ . With this choice,  $f(x)$  has the minimum value

$$f(x) \geq f(x^*) + \left( \frac{1}{2} \alpha^2 d^T \nabla^2 f(x^*) d - |o(\alpha^2)| \right)$$

which satisfies  $f(x^*) < f(x)$  by virtue of the positive nature of the parenthesized term. And so, with  $\varepsilon < \alpha$ ,  $x^*$  is necessarily a local minimum of  $f$ .

- 1.2. Give an example where a local minimum  $x^*$  is not a regular point and the above necessary condition is false (be sure to justify both of these claims).

Consider the system

$$f(x, y, z) = (x - 1)^2 + (y - 1)^2 + z^2$$

$$h_1(x, y, z) = z$$

$$h_2(x, y, z) = x^2 + y^2 + (z - 1)^2 - 1$$

The point  $(0, 0, 0)$  is trivially a local minimum of this system (it is the only point that lies in  $D$ ) yet the condition (1.24) is false. This is because the gradients of the above functions at  $(0, 0, 0)$  are

$$\nabla f(0, 0, 0) = \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix}$$

$$\nabla h_1(0, 0, 0) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\nabla h_2(0, 0, 0) = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$$

and so no solution exists for the equation

$$\nabla f(0, 0, 0) + \lambda_1 \nabla h_1(0, 0, 0) + \lambda_2 \nabla h_2(0, 0, 0) = 0.$$

- 1.3. Generalize the previous argument to an arbitrary number  $m \geq 1$  of equality constraints (still assuming that  $x^*$  is a regular point).

Assume that  $x^*$  is a local minimum of  $f$  over  $D$  where  $D$  is defined by  $m$  equality constraints. Consider  $m + 1$  arbitrary vectors  $d_1, \dots, d_{m+1} \in \mathbb{R}^n$  and the following map,

$$F : (\alpha_1, \dots, \alpha_{m+1}) \rightarrow (f(x^* + \sum_{k=1}^{m+1} \alpha_k d_k), h_1(x^* + \sum_{k=1}^{m+1} \alpha_k d_k), \dots, h_m(x^* + \sum_{k=1}^{m+1} \alpha_k d_k)).$$

The Jacobian of this map is given by

$$\begin{pmatrix} \nabla f(x^*) \cdot d_1 & \cdots & \nabla f(x^*) \cdot d_{m+1} \\ \nabla h_1(x^*) \cdot d_1 & \cdots & \nabla h_1(x^*) \cdot d_{m+1} \\ \vdots & \ddots & \vdots \\ \nabla h_m(x^*) \cdot d_1 & \cdots & \nabla h_m(x^*) \cdot d_{m+1} \end{pmatrix}$$

and, using the same argument as in the book, it must be singular. In brief: if it were not, the inverse function theorem guarantees we could find a point satisfying the constraints where  $f$  attains a smaller value than  $f(x^*)$ —but this would contradict the fact that  $x^*$  is a local minimum.

Choose each  $d_k$  such that  $\nabla h_k(x^*) \cdot d_k \neq 0$  for  $k = 1, \dots, m$  (this is possible because every  $\nabla h_k(x^*)$  is nonzero by the condition of regularity). Since the matrix is singular and every row except the first contains a nonzero element, it must be that the first row is some linear combination of the remaining rows, i.e. each entry in the top row takes the form

$$\nabla f(x^*) \cdot d_k = -(\lambda_1^* \nabla h_1(x^*) + \cdots + \lambda_m^* \nabla h_m(x^*)) \cdot d_k.$$

In particular, since  $d_{m+1}$  can be arbitrarily chosen, this must hold for all possible  $d_{m+1}$ , implying

$$\nabla f(x^*) + \lambda_1^* \nabla h_1(x^*) + \cdots + \lambda_m^* \nabla h_m(x^*) = 0.$$

- 1.4. Consider a curve  $D$  in the plane described by the equation  $h(x) = 0$ , where  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$  is a  $C^1$  function. Let  $y$  and  $z$  be two fixed points in the plane, lying on the same side with respect to  $D$  (but not on  $D$  itself). Suppose that a ray of light emanates from  $y$ , gets reflected off  $D$  at some point  $x^* \in D$ , and arrives at  $z$ . Consider the following two statements: (i)  $x^*$  must be such that the total Euclidean distance traveled by light to go from  $y$  to  $z$  is minimized over all nearby candidate reflection points  $x \in D$  (Fermat's principle); (ii) the angles that the light ray makes with the line normal to  $D$  at  $x^*$  before and after the reflection must be the same (the law of reflection). Accepting the first statement as a hypothesis, prove that the second statement follows from it, with the help of the first-order necessary condition for constrained optimality.

The function we need to minimize is the Euclidean distance between  $y$  and  $z$  via reflection about  $x$ ,

$$f(x) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} + \sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2}.$$

The gradient of this function is

$$\nabla f(x) = \frac{1}{\sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2}} \begin{bmatrix} x_1 - z_1 \\ x_2 - z_2 \end{bmatrix} + \frac{1}{\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix}$$

and we can see that it consists of two vectors: the normalized vector from  $z$  to  $x$ , hereby labeled  $\hat{\mathbf{v}}_{zx}$ , and the normalized vector from  $y$  to  $x$ ,  $\hat{\mathbf{v}}_{yx}$ .

Using the first-order necessary condition for constrained optimality, we see that this must be proportional to the gradient of  $h(x)$ , i.e.  $\nabla f(x) = \hat{\mathbf{v}}_{zx} + \hat{\mathbf{v}}_{yx}$  is normal to the line described by  $h(x)$ . Geometrically, the only way this can occur is if the normalized vectors flank  $\nabla h(x)$  at equal angles.

- 1.5. Consider the space  $V = C^0([0, 1], \mathbb{R})$ , let  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  be a  $C^1$  function, and define the functional  $J$  on  $V$  by  $J(y) = \int_0^1 \phi(y(x)) dx$ . Show that its first variation exists and is given by the formula  $\delta J|_y(\eta) = \int_0^1 \varphi'(y(x))\eta(x) dx$ .

Let us start with (1.33),

$$\delta J|_y(\eta) = \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} \left( \int_0^1 \varphi(y(x) + \alpha\eta(x)) dx - \int_0^1 \varphi(y(x)) dx \right)$$

As  $\alpha$  approaches 0, we may replace  $\varphi(y(x) + \alpha\eta(x))$  with its first order linear approximation about  $y(x)$  (which is possible because  $\varphi$  is once differentiable). In particular, we have

$$\begin{aligned} \delta J|_y(\eta) &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} \left( \int_0^1 \varphi(y(x)) + \alpha\varphi'(y(x))\eta(x) dx - \int_0^1 \varphi(y(x)) dx \right) \\ &= \lim_{\alpha \rightarrow 0} \frac{\alpha}{\alpha} \int_0^1 \varphi'(y(x))\eta(x) dx \\ &= \int_0^1 \varphi'(y(x))\eta(x) dx \end{aligned}$$

- 1.6. Consider the same functional  $J$  as in Exercise 1.5, but assume now that  $\psi$  is  $C^2$ . Derive a formula for the second variation of  $J$  (make sure that it is indeed a quadratic form).

The second variation is given by the functional derivative of  $\delta J|_y(\eta)$

$$\begin{aligned} \delta^2 J|_y(\eta, \xi) &= \lim_{\alpha \rightarrow 0} \frac{\delta J|_{y+\alpha\xi}(\eta) - \delta J|_y(\eta)}{\alpha} \\ &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} \left( \int_0^1 \varphi'(y(x) + \alpha\xi(x))\eta(x) dx - \int_0^1 \varphi'(y(x))\eta(x) dx \right) \end{aligned}$$

where we have used  $\xi$  to aid in the expansion process with the understanding that it will eventually be set to  $\eta$ . Since  $\varphi(y(x))$  is twice differentiable, we may replace  $\varphi'(y(x))$  with its first order linear approximation to find

$$\begin{aligned} \delta^2 J|_y(\eta, \xi) &= \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} \left( \int_0^1 \varphi'(y(x))\eta(x) + \alpha\varphi''(y(x))\xi(x)\eta(x) dx - \int_0^1 \varphi'(y(x))\eta(x) dx \right) \\ &= \lim_{\alpha \rightarrow 0} \frac{\alpha}{\alpha} \int_0^1 \varphi''(y(x))\xi(x)\eta(x) dx \\ &= \int_0^1 \varphi''(y(x))\xi(x)\eta(x) dx \end{aligned}$$

Replacing  $\xi$  with  $\eta$  gives us the second variation of  $J$  with respect to  $y$ . This is a bilinear functional with both of its arguments given by  $\eta$ , i.e. a quadratic form.

- 1.7. Give an example of a function space  $V$ , a norm on  $V$ , a closed and bounded subset  $A$  of  $V$ , and a continuous functional  $J$  on  $V$  such that a global minimum of  $J$  over  $A$  does not exist. (Be sure to demonstrate that all the requested properties hold.)