Oregon State University

Homework 2. Basic properties of solutions

Hyuntaek Oh

ohhyun@oregonstate.edu

Due: April 23, 2025

ECE599/ AI539 Nonlinear Optimization (Spring 2025)
Homework 1. Preliminaries (Due: 11:59pm, Wednesday, April 16)

Instruction: Students should provide enough detail of the logical procedure of deriving answers. Answers without sufficient justification will receive partial or no credit.

Reading: Before working on the homework questions, read Sections 7.1-7.5 and Section 7.7 of the textbook (Luenberger and Ye).

1. Exercise 1 of Chapter 7. (Hint: Reviewing Example 2 of Section 7.2 may help)

To approximate a function g over the interval $[0, 1]$ by a polynomial p of degree n (or less), we minimize the criterion

$$f(\mathbf{a}) = \int_0^1 [g(x) - p(x)]^2 \, dx,$$

where $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$. Find the equations satisfied by the optimal coefficients $\mathbf{a} = (a_0, a_1, \ldots, a_n)$.

According to Example 2 in chapter 7.2, $f(\mathbf{a})$ is:

$$f(\mathbf{a}) = \int_0^1 [g(x) - p(x)]^2 dx = \int_0^1 g(x)^2 - 2g(x)p(x) + p(x)^2 dx$$

$$\Rightarrow \int_0^1 g(x)^2 dx - 2\sum_{j=0}^{n} a_j \int_0^1 g(x)x^j dx + \sum_{i=0}^{n}\sum_{j=0}^{n} a_i a_j \int_0^1 x^{i+j} dx$$

$$\Rightarrow \mathbf{a}^T Q\mathbf{a} - 2\mathbf{b}^T\mathbf{a} + c$$

, where

$$Q_{ij} = \int_0^1 x^{i+j} dx = \frac{x^{i+j+1}}{i+j+1} = \frac{1}{i+j+1}$$

$$b_i = \int_0^1 g(x)x^i dx$$

$$c = \int_0^1 g(x)^2 dx, \text{ c is negligible since it is constant}$$

For optimization, $f(\mathbf{a})$ needs to be zero, so it would be:

$$\nabla f(\mathbf{a}) = 0$$

$$\Rightarrow \nabla f(\mathbf{a}) = 2Q\mathbf{a} - 2\mathbf{b}$$
$$\Rightarrow 2Q\mathbf{a} - 2\mathbf{b} = 0$$
$$Q\mathbf{a} = \mathbf{b}$$

From this result, the equations would be:

$$\sum_{j=0}^{n} \int_0^1 x^{i+j} dx \cdot a_j = \int_0^1 g(x) x^i dx, \text{ where } i = 0, 1, ..., n$$

$$\Rightarrow \sum_{j=0}^{n} \frac{a_j}{i+j+1} = \int_0^1 g(x) x^i dx$$

This means that loss function $[g(x) - p(x)]$ *is orthogonal to every* x^i.

2. Exercise 2 of Chapter 7. (Assume that the solution also satisfies $x_2 > 0$)

In Example 3 of Section 7.2 show that if the solution has $x_1 > 0$, $x_1 + x_2 = 1$, then it is necessary that

$$b_1 - b_2 + (c_1 - c_2) h(x_1) = 0$$
$$b_2 + (c_2 - c_3) h(x_1 + x_2) \leqslant 0.$$

Hint: One way is to reformulate the problem in terms of the variables x_1 and $y = x_1 + x_2$.

To prove the given equations, the problem is reformulated in terms of the variables x_1 *and* $y = x_1 + x_2$, *meaning that:*

$$x_1, \quad y := x_1 + x_2$$

The constraints in Example 3 of Section 7.2 are:

$$x_1 \geq 0$$

$$x_2 \geq 0 \Rightarrow y - x_1 \geq 0 \Rightarrow y \geq x_1$$
$$x_1 + x_2 \leq 1 \Rightarrow y \leq 1$$

The reformulated cost function is:

$$f(x, y) = b_1 x + b_2(y - x) + c_1 \int_0^x h(t) dt + c_2 \int_x^y h(t) dt + c_3 \int_y^1 h(t) dt$$

$$= (b_1 - b_2)x + b_2 y + c_1 \int_0^x h(t)dt + c_2 \int_x^y h(t)dt + c_3 \int_y^1 h(t)dt$$

These constraints becomes:

$$0 \le x \le y \le 1$$

Based on the new constraints and reformulated function, the objective function is:

$$\min_{0 \le x \le y \le 1} f(x, y)$$

In the problem statement, however, the solution has $x_1 > 0$, $x_2 > 0$, and $x_1 + x_2 = y = 1$, meaning that the lower bound $x_1 = 0$ is not active, and the upper bound $y \le 1$ is active. So, the feasible directions are any directions under the constraints.

According to the first-order necessary condition, If $\mathbf{x}^*$ is a local minimum point of f over Ω, then for any $\mathbf{d} \in \mathbb{R}^n$ that is a feasible direction at $\mathbf{x}^*$, we have $\nabla f(\mathbf{x}^*)d \ge 0$.

Based on the first-order necessary condition, for all $[d_1, d_2]^T$ multiplying with any gradient such as $\nabla g(x_1, y)$, the inequality satisfies greater than or equal to zero:

$$\nabla g(x_1, y)[d_1, d_2]^T \ge 0$$

To satisfy the condition, d_1 and d_2, the elements of the direction vector, are any real value and less than and equal to zero respectively:

$$d_1 = \text{any real value}$$

$$d_2 \le 0$$

Moreover, the gradient $\nabla g(x_1, y)$ vector defines that opposes all directions in which movement is allowed (feasible). If $\nabla f(\mathbf{x})^T d < 0$, moving in direction d would decrease the function value. If $\nabla f(\mathbf{x})^T d = 0$, the gradient is orthogonal to d, and movement along d does not change the objective.

Additionally, the function f can be differentiate with respect to x and y.

The derivative with respect to x would be:

$$f(x, y) = (b_1 - b_2)x + b_2 y + c_1[h(x) - h(0)] + c_2[h(y) - h(x)] + c_3[h(1) - h(y)]$$

$$\frac{\partial f}{\partial x} = b_1 - b_2 + c_1 h(x) - c_2 h(x)$$

Since $x > 0$, the constraint $x \ge 0$ is not active, so the derivative must be 0 at the optimum:

$$(b_1 - b_2) + (c_1 - c_2)h(x) = 0$$

Oregon State University
College of Engineering

$$\Rightarrow (b_1 - b_2) + (c_1 - c_2)h(x_1) = 0$$

On the other hand, the derivative with respect to y would be:

$$\frac{\partial f}{\partial y} = b_2 + c_2 h(y) - c_3 h(y)$$

To find out a minimum point, the function f needs to increase or stay when y is reduced, so it would be:

$$b_2 + c_2 h(y) - c_3 h(y) \leq 0$$
$$b_2 + (c_2 - c_3)h(y) \leq 0$$
$$\Rightarrow b_2 + (c_2 - c_3)h(x_1 + x_2) \leq 0$$

Thus, the solution holds.

3. Exercise 3 of Chapter 7.

   (a) Using the first-order necessary conditions, find a minimum point of the function

$$f(x, y, z) = 2x^2 + xy + y^2 + yz + z^2 - 6x - 7y - 8z + 9$$

By using the first-order necessary conditions, which means that the gradient of a function is equal to zero, the function above would be:

$$\frac{\partial f(x,y,z)}{\partial x} = 0 \Rightarrow 4x + y - 6 = 0 \tag{1}$$

$$\frac{\partial f(x,y,z)}{\partial y} = 0 \Rightarrow y + 2z - 8 = 0 \tag{2}$$

$$\frac{\partial f(x,y,z)}{\partial z} = 0 \Rightarrow x + 2y + z - 7 = 0 \tag{3}$$

From (1) and (2), $y = -4x + 6$ and $y = -2z + 8$ can be used to get the relation between x and z:

$$-4x + 6 = -2z + 8 \Rightarrow x = \frac{1}{2}z - \frac{1}{2}$$

That result and (2) can be applied to (3) to get z:

$$\frac{1}{2}z - \frac{1}{2} - 4z + 16 + z - 7 = 0$$

$$z = \frac{17}{5}$$

By using z, y in (2) would be:

$$y = -2 * \frac{17}{5} + 8 = \frac{6}{5}$$

Finally, x in (1) would be:

$$\frac{6}{5} = -4 * x + 6 \Rightarrow x = \frac{6}{5}$$

Thus, a minimum point of the function $f(x, y, z)$ is $(x^*, y^*, z^*) = (\frac{6}{5}, \frac{6}{5}, \frac{17}{5})$

(b) Verify that the point is a relative minimum point by verifying that the second-order sufficiency conditions hold.

The second-order sufficiency conditions are:

i. $\nabla f(\mathbf{x}^*) = 0$

ii. $\mathbf{F}(\mathbf{x}^*)$ *is positive definite*

From the problem (a), a minimum point of the function, which means that $\nabla f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$ is found, $(x^*, y^*, z^*) = (\frac{6}{5}, \frac{6}{5}, \frac{17}{5})$.

To check the resulted point is a relative minimum point, Hessian matrix of f $F(\mathbf{x}^*)$ or $\nabla^2 f$ can be used. According to Appendix part A in the textbook, this is:

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Then, if all eigenvalue $\lambda_i > 0$ in the Hessian matrix above, it is positive definite, meaning that:

$$det(H - \lambda I) = 0$$

$$\Rightarrow det(\begin{bmatrix} 4-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 2-\lambda \end{bmatrix}) = 0$$

$$\Rightarrow (4-\lambda)[(2-\lambda)^2 - 1] - (2-\lambda)$$

By extending the mathematical expression above:

$$-\lambda^3 + 8\lambda^2 - 18\lambda + 10 = 0 \text{ or } \lambda^3 - 8\lambda^2 + 18\lambda - 10 = 0$$

The possible eigenvalues would be:

$$\lambda_1 \approx 1.46, \quad \lambda_2 \approx 2.38, \quad \lambda_3 \approx 4.16$$

This means all eigenvalues are positive, and then the Hessian matrix is positive definite. Thus, the second-order sufficiency conditions holds, meaning that the point $(x^*, y^*, z^*) = (\frac{6}{5}, \frac{6}{5}, \frac{17}{5})$ *is a relative minimum point (local minimum).*

(c) Prove that the point is a global minimum point.

If the function f *is convex and* $\nabla f(\boldsymbol{x}^*) = 0$*, then* $\boldsymbol{x}^*$ *is a global minimum.*

To verify that the function $f(x, y, z)$ *is convex, the Hessian matrix of* f *is positive semidefinite for all* x*. However, from the problem 2 above, the Hessian matrix of* f *is positive definite, meaning that the function is strictly convex.*

Furthermore, the point (x^*, y^*, z^*) *satisfies the first-order condition,* $\nabla f(x^*, y^*, z^*) = 0$*, it follows by the second-order sufficiency theorem for convex functions that this point is the unique global minimizer of* f*.*

Thus, the optimal point (x^*, y^*, z^*) *is a global minimum point.*