Numerical Optimization Assignment 2

David Bulovic

December 2021

1 Lagrange Multiplier Problem

a) min $x_2 - x_1$ s.t. $x_1 \geq 4x_2$, $x_2 = \frac{1}{10}x_1^2 - 3$ We can begin by formulating the equation:

$$\mathcal{L}(x_1, x_2, \lambda, \mu) = x_2 - x_1 + \lambda(\frac{1}{10}x_1^2 - x_2 - 3) + \mu(4x_2 - x_1)$$

The derivatives are:

$$\nabla_{x_1}\mathcal{L} = -1 + \lambda\frac{2}{10}x_1 - \mu \stackrel{!}{=} 0 \tag{1}$$

$$\nabla_{x_2}\mathcal{L} = 1 - \lambda + 4\mu \stackrel{!}{=} 0 \tag{2}$$

$$\nabla_{\lambda}\mathcal{L} = \frac{1}{10}x_1^2 - x_2 - 3 \stackrel{!}{=} 0 \tag{3}$$

$$\mu(g(x)) = 0 \tag{4}$$

Case 1: $\mu > 0$, $g_1 = 0$(inequality constraint is binding)

$$4x_2 - x_1 = 0$$

$$x_1 = 4x_2 \tag{5}$$

Using this and the equation (3):

$$\frac{16}{10}x_2^2 - x_2 - 3 = 0$$

$$x_2 = \frac{1 \pm \sqrt{1 + 12 * \frac{16}{10}}}{\frac{16}{5}}$$

$$x_2 = \frac{5}{16} \pm \frac{\sqrt{505}}{16}$$

Inserting back into (5):

$$x_1 = \frac{5}{4} \pm \frac{\sqrt{505}}{4}$$

This means that our first candidate points are:

$$
\begin{aligned}
(x_1, x_2) &= (\frac{5}{4} + \frac{\sqrt{505}}{4}, \frac{5}{16} + \frac{\sqrt{505}}{16}) \\
&= (\frac{5}{4} - \frac{\sqrt{505}}{4}, \frac{5}{16} - \frac{\sqrt{505}}{16})
\end{aligned}
$$

Case 2: $\mu = 0$, g not binding From (1) and (2) we see:

$$-1 + \lambda \frac{2}{10} x_1 = 0$$

$$1 - \lambda = 0 \implies \lambda = 1$$

$$-1 + \frac{2}{10} x_1 = 0 \implies x_1 = 5$$

Using (3):

$$
\begin{aligned}
\frac{1}{10} 25 - x_2 - 3 = 0 \\
x_2 = \frac{25}{10} - 3 \\
x_2 = -\frac{1}{2}
\end{aligned}
$$

Next candidate point: $(x_1, x_2) = (5, -\frac{1}{2})$. Calculating the values of the function for every point:

$$
\begin{aligned}
f(\frac{5}{4} + \frac{\sqrt{505}}{4}, \frac{5}{16} + \frac{\sqrt{505}}{16}) &\approx -5.15 \\
f(\frac{5}{4} - \frac{\sqrt{505}}{4}, \frac{5}{16} - \frac{\sqrt{505}}{16}) &\approx 3.28 \\
f(5, -\frac{1}{2}) &= -5.5
\end{aligned}
$$

From this we can conclude that the point $(5, -\frac{1}{2})$ is the optimal solution. In the file *figures.pdf*, this point is marked with a black spot, while the other candidates are red.

b) min $||x||_2^2$ s.t. $x_2 \geq 3 - x_1$, $x_2 \geq 2$. We can reformulate:

$$
\begin{aligned}
f(\mathbf{x}) &= x_1^2 + x_2^2 \\
g_1(\mathbf{x}) &= 3 - x_1 - x_2 \\
g_2(\mathbf{x}) &= 2 - x_2
\end{aligned}
$$

$$\mathcal{L}(x_1, x_2, \mu_1, \mu_2) = x_1^2 + x_2^2 + \mu_1(3 - x_1 - x_2) + \mu_2(2 - x_2)$$

$$\nabla_{x_1} \mathcal{L} = 2x_1 - \mu_1 \overset{!}{=} 0 \tag{6}$$

$$\nabla_{x_2} \mathcal{L} = 2x_2 - \mu_1 - \mu_2 \overset{!}{=} 0 \tag{7}$$

$$\mu_1 g_1(\mathbf{x}) = 0 \tag{8}$$

$$\mu_2 g_2(\mathbf{x}) = 0 \tag{9}$$

Case 1: $\mu_1 > 0$ and $\mu_2 > 0$ (both binding). From (8) and (9):

$$3 - x_1 - x_2 = 0$$

$$2 - x_2 = 0 \implies x_2 = 2$$
$$\implies x_1 = 1$$

First candidate point: $(1, 2)$
Case 2: $\mu_1 = 0$ and $\mu_2 > 0$

$$2 - x_2 = 0 \implies x_2 = 2$$

From (6): $2x_1 = 0 \implies x_1 = 0$. Since $2 < 3 - 0$, first constraint is not satisfied, therefore not a solution.
Case 3: $\mu_1 > 0$ and $\mu_2 = 0$

$$3 - x_1 - x_2 = 0 \implies x_2 = 3 - x_1$$

$$\begin{aligned} \text{From (6): } & 2x_1 - \mu_1 = 0 \\ \text{From (7): } & 2(3 - x_1) - \mu_1 = 0 \\ & 6 - 2x_1 - \mu_1 = 0 \end{aligned}$$

Adding the two equations above:

$$6 - 2\mu_1 = 0$$

$$\mu_1 = 3 \implies x_1 = \frac{3}{2} \implies x_2 = \frac{3}{2}$$

Since $\frac{3}{2}$ is smaller than 2, this is not a valid candidate.
Case 4: $\mu_1 = 0$ and $\mu_2 = 0$
From (6) and (7):

$$\begin{aligned} 2x_1 &= 0 \\ 2x_2 &= 0 \\ \implies (x_1, x_2) &= (0, 0) \end{aligned}$$

Since the second constraint is: $x_2 \geq 2$, this is clearly not a solution.
We can conclude that the point $(x_1, x_2) = (1, 2)$ is the optimum, with the value 5.

c) min $(x_1 - 1)^2 + x_1 x_2^2 - 2$ s.t. $x_1^2 + x_2^2 \leq 4$

$$\mathcal{L}(x_1, x_2, \mu) = x_1^2 - 2x_1 + 1 + x_1 x_2^2 - 2 + \mu(x_1^2 + x_2^2 - 4)$$

$$\nabla_{x_1}\mathcal{L} = 2x_1 - 2 + x_2^2 + 2\mu x_1 \stackrel{!}{=} 0 \tag{10}$$

$$\nabla_{x_2}\mathcal{L} = 2x_1x_2 + 2\mu x_2 \overset{!}{=} 0 \tag{11}$$

$$\mu g_1(\mathbf{x}) = 0 \tag{12}$$

Case 1: $\mu > 0$

$$x_1^2 + x_2^2 - 4 = 0$$

Using this and equations (10) and (11), we get a system of three equations, the solutions for which are (for x_1 and x_2):

$$\begin{aligned}
(x_1, x_2) &= (-2, 0) \\
&= (2, 0) \\
&= \left(\frac{1}{3}(1-\sqrt{7}), -\frac{1}{3}\sqrt{2(14+\sqrt{7})}\right) \\
&= \left(\frac{1}{3}(1-\sqrt{7}), \frac{1}{3}\sqrt{2(14+\sqrt{7})}\right) \\
&= \left(\frac{1}{3}(1+\sqrt{7}), -\frac{1}{3}\sqrt{2(14-\sqrt{7})}\right)
\end{aligned}$$

Case 2: $\mu = 0$

From (10) and (11):

$$2x_1 - 2 + x_2^2 = 0$$

$$2x_1x_2 = 0$$

If $x_1 = 0$:

$$\begin{aligned}
-2 + x_2^2 &= 0 \\
x_2 &= \pm\sqrt{2}
\end{aligned}$$

Points: $(0, \pm\sqrt{2})$.

If $x_2 = 0$:

$$\begin{aligned}
2x_1 - 2 &= 0 \\
x_1 &= 1
\end{aligned}$$

Second point: $(1, 0)$. Calculating the values of the function at the candidate points:

$$f(-2, 0) = 7$$

$$f(2, 0) = -1$$

$$f\left(\frac{1}{3}(1-\sqrt{7}), -\frac{1}{3}\sqrt{2(14+\sqrt{7})}\right) \approx -1.63$$

$$f\left(\frac{1}{3}(1-\sqrt{7}), \frac{1}{3}\sqrt{2(14+\sqrt{7})}\right) \approx -1.63$$

$$f\left(\frac{1}{3}(1+\sqrt{7}), -\frac{1}{3}\sqrt{2(14-\sqrt{7})}\right) \approx 1.11$$
$$f(0,\sqrt{2}) = (0-1)^2 + 0(\sqrt{2})^2 - 2 = -1$$
$$f(0,-\sqrt{2}) = (0-1)^2 + 0(-\sqrt{2})^2 - 2 = -1$$
$$f(1,0) = (1-1)^2 + 1 * 0^2 - 2 = -2$$

The point $(1,0)$ is the optimum.

2 Lagrange Augmentation

1.
$$\min\ f(x) = (x_1-1)^2 - x_1x_2 \quad \text{s.t.} \quad -x_1+4 = x_2$$

Equality constraint: $h(x) = x_1 + x_2 - 4$

$$\mathcal{L}(x_1,x_2,\lambda) = x_1^2 - 2x_1 + 1 - x_1x_2 + \lambda(x_1+x_2-4)$$

$$\nabla_{x_1}\mathcal{L} = 2x_1 - 2 - x_2 + \lambda \stackrel{!}{=} 0 \tag{13}$$
$$\nabla_{x_2}\mathcal{L} = -x_1 + \lambda \stackrel{!}{=} 0 \tag{14}$$
$$\nabla_{\lambda}\mathcal{L} = x_1 + x_2 - 4 \stackrel{!}{=} 0 \tag{15}$$

From (14) we see that: $\lambda = x_1$. Inserting this into (13):

$$2x_1 - 2 - x_2 + x_1 = 0$$

$$3x_1 - x_2 = 2$$

Adding this to (15) gives us:

$$3x_1 - x_2 + x_1 + x_2 = 2 + 4$$

$$4x_1 = 6$$
$$x_1 = \frac{3}{2} \implies \lambda = \frac{3}{2}$$

Putting $x_1 = \frac{3}{2}$ into (15):
$$\frac{3}{2} + x_2 = 4$$

$$x_2 = \frac{5}{2}$$

The solution is $(x_1,x_2) = (\frac{3}{2},\frac{5}{2})$, and $\lambda = \frac{3}{2}$.

2.

$$\min_{x_k}\{\max_{\lambda_k} f(x_k) + \lambda_k^T h(x_k) - \frac{1}{2\alpha}||\lambda_k - \lambda_{k-1}||^2 = \mathcal{L}_a(x_k,\lambda_k,\alpha)\}$$

To solve the inner optimization problem, we can derive for λ_k:
(Avoided using the sum notation for simpler calculation, since there is only one equality constraint)

$$\nabla_{\lambda_k}\mathcal{L}_a = h(x_k) - \frac{1}{\alpha}(\lambda_k - \lambda_{k-1}) =^! 0$$

$$\frac{1}{\alpha}(\lambda_k - \lambda_{k-1}) = h(x_k)$$

Here we get an update rule for λ_k:

$$\lambda_k = \alpha h(x_k) + \lambda_{k-1}$$

In our example this would be:

$$\lambda_k = \alpha(-x_1 - x_2 - 4) + \lambda_{k-1}$$

Plugging this into the initial equation:

$$\min_{x_k} f(x_k) + (\alpha h(x_k) + \lambda_{k-1})h(x_k) - \frac{1}{2\alpha}||(\alpha h(x_k) + \lambda_{k-1}) - \lambda_{k-1}||^2$$

$$\min_{x_k} f(x_k) + \alpha||h(x_k)||^2 + \lambda_{k-1}h(x_k) - \frac{\alpha}{2}||h(x_k)||^2$$

$$\min_{x_k} f(x_k) + \lambda_{k-1}h(x_k) + \frac{\alpha}{2}||h(x_k)||^2$$

Which is what we had to prove.

3.

$$\min_{x_k} f(x_k) + \lambda_{k-1}h(x_k) + \frac{\alpha}{2}||h(x_k)||^2$$

Substituting for our problem:

$$\min_{x_k} x_1^2 - 2x_1 + 1 - x_1x_2 + \lambda_{k-1}(-x_1 - x_2 + 4) + \frac{\alpha}{2}(-x_1 - x_2 + 4)^2$$

Deriving by x_k:

$$\nabla_{x_1} = 2x_1 - 2 - x_2 - \lambda_{k-1} + \alpha(-x_1 - x_2 + 4)(-1) =^! 0$$

$$\nabla_{x_2} = -x_1 - \lambda_{k-1} + \alpha(-x_1 - x_2 + 4)(-1) =^! 0$$

This gives us a system of linear equations, which we can simply solve for x_1 and x_2 in terms of λ_{k-1} and α to get the update rules for x_k:

$$x_1 = \frac{6\alpha + \lambda_{k-1}}{4\alpha - 1}$$

$$x_2 = \frac{10\alpha + 3\lambda_{k-1} + 2}{4\alpha - 1}$$

4. - 8.
The contour plot with the constraints is available in the *figures.pdf* file. Also the calculated optimal solution is presented with a black spot. The implemented iterative algorithm is presented with white crosses, with $\lambda_{k-1} = -1$ and $\alpha = 0.7$.

9.
Increasing the value of α seems to make the iterations of x_k more dense. Also, the iterations closest to the actual solution seem to be more precise for larger values of α.

3 Least Squares Fitting

1. The data is plotted in the *figures.pdf* file. Due to the general shape of the scatter plot it can be determined that $d = 3$.

2. The linear system was formulated in the python file *main.py*.

$$\mathbf{Ax} \approx \mathbf{b}$$

Where the rows of the matrix A correspond to the double sum of the x_n and y_n multiplications, x are the coefficients, and b are the values z

3. Calculating the LS solution with the formula:

$$x = (A^T A)^{-1} A^T b$$

We get the the following values for the coefficients (up to 4 significant digits):

$$\begin{aligned}
(0.9593,&\quad 0.9060,\quad 0.5267,\quad -0.0007,\\
0.0412,&\quad 0.0496,\quad 0.8042,\quad -0.0148,\\
0.4926,&\quad -0.0300,\quad 0.0004,\quad 0.0016,\\
-0.4687,&\quad 0.0058,\quad -0.0019,\quad 0.0006)
\end{aligned}$$

The corresponding wireplot is also presented in the *figures.pdf* file.