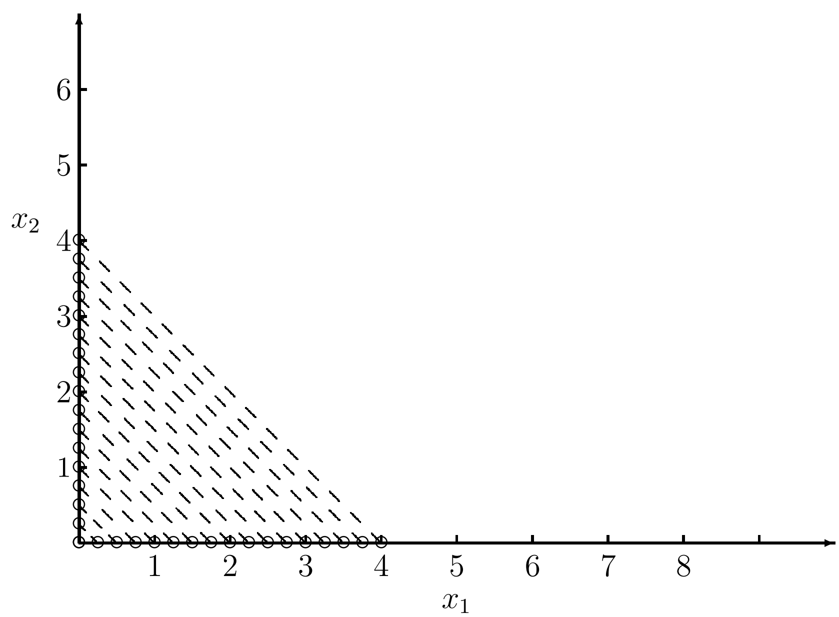
Algorithms for Optimal Decisions
Tutorial 4
Answers

Exercise 1

Solve the following **Q.P.** *using the Frank–Wolfe method:*

$$\begin{aligned}\min_x f(x) &= x_1^2 - x_1x_2 + x_2^2 - 3x_1 \\ s.t. \quad & -x_1 \le 0 \\ & -x_2 \le 0 \\ & x_1 + x_2 - 4 \le 0.\end{aligned} \tag{1}$$

Starting point : $x^{(0)} = (x_1^{(0)}, x_2^{(0)}) = (0,0)$.



Solution :

It is easy to see that the feasible region is a triangle with vertices $\{(0,0),(4,0),(0,4)\}$. Since **LP** solutions are always at vertices of the feasible region, it will be easy to solve any sub–**LP** problem by testing at each vertex. The gradient of the objective function is given by:

$$\nabla f(x) = \begin{pmatrix} 2x_1 - x_2 - 3 \\ -x_1 + 2x_2 \end{pmatrix} \tag{2}$$

We start with $x^{(0)} = (0,0)$:

$$f(x^{(0)}) = 0, \quad \nabla f(x^{(0)}) = \begin{pmatrix} -3 \\ 0 \end{pmatrix}. \tag{3}$$

We will minimize

$$\begin{array}{ll} \min\limits_{x} & -3x_1 + 0x_2 \\ s.t. & x_1 + x_2 \leq 4 \\ & x_1 \geq 0, \quad x_2 \geq 0, \end{array}$$

to find the furthest in that direction we can go. Clearly, the optimal vertex is

$$x_{LP}^{(1)} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}. \tag{4}$$

The new point will be

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \tau \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 4\tau \\ 0 \end{pmatrix}. \tag{5}$$

$$f(x^{(1)}) = (4\tau)^2 - 3(4\tau), \quad \frac{d}{d\tau} f(x^{(1)}) = 32\tau - 12 = 0 \Rightarrow \tau = \frac{3}{8}. \tag{6}$$

Replacing $\tau = \frac{3}{8}$ into (5) we obtain:

$$x^{(1)} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix}, \quad f(x^{(1)}) = -2.25. \tag{7}$$

We now start a new iteration:

$$\nabla f(x^{(1)}) = \begin{pmatrix} 0 \\ -\frac{3}{2} \end{pmatrix}. \tag{8}$$

We will minimize

$$\begin{array}{ll} \min\limits_{x} & 0x_1 - \dfrac{3}{2}x_2 \\ s.t. & x_1 + x_2 \leq 4 \\ & x_1 \geq 0, \quad x_2 \geq 0, \end{array}$$

to find the furthest in that direction we can go. Clearly, the optimal vertex is

$$x_{LP}^{(2)} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}. \tag{9}$$

The new point will be

$$x^{(2)} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} + \tau \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} - \frac{3}{2}\tau \\ 4\tau \end{pmatrix}. \tag{10}$$

$$f(x^{(2)}) = -\frac{9}{4} - 6\tau + \frac{97}{4}\tau^2, \quad \frac{d}{d\tau} f(x^{(2)}) = \frac{97}{2}\tau - 6 = 0 \Rightarrow \tau = \frac{12}{97}. \tag{11}$$

Replacing $\tau = \frac{12}{97}$ into (10) we obtain:

$$x^{(2)} = \begin{pmatrix} 1.314 \\ 0.496 \end{pmatrix}, \quad f(x^{(2)}) = -2.621. \tag{12}$$

We now start a new iteration...

The new point will be

$$x^{(3)} = \begin{pmatrix} 1.314 + 2.686\tau \\ 0.496 - 0.496\tau \end{pmatrix}, \tag{13}$$

$$\begin{aligned} f(x^{(3)}) &= -2.621 - 2.172\tau + 8.793\tau^2, \\ \frac{d}{d\tau} f(x^{(3)}) &= -2.172 + 17.586\tau = 0 \Rightarrow \tau = 0.124. \end{aligned} \tag{14}$$

Next point is:

$$x^{(3)} = \begin{pmatrix} 1.647 \\ 0.434 \end{pmatrix}, \quad f(x^{(3)}) = -2.755. \tag{15}$$

This is getting cumbersome to continue by hand, so we shall stop here. Even though $x^{(3)}$ is not far from $x^{(2)}$, $\nabla f(x^{(2)})$ is not close to $(0,0)^t$ and we have ways to go before approaching convergence.

We can easily find the unconstrained optimum $x^* = (2,1)^t$ by solving

$$\nabla f(x) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{16}$$

Armijo Stepsize Rule

Fix η and γ to fractional values (i.e. $\eta = \gamma = \frac{1}{2}$). Test for the sequence $\tau = \gamma^0, \gamma^1, \gamma^2 ...$ until the improvement $f(x_{k+1}) - f(x_k)$ is better than a certain amount related to the gradient:

$$f(x_{k+1}) - f(x_k) \leq -\eta\tau(\nabla f(x_k)^t \nabla f(x_k)). \tag{17}$$

Example :

$$x^{(1)} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix}; \quad f(x^{(1)}) = -\frac{9}{4}; \quad x^{(2)} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} + \tau \begin{pmatrix} -\frac{3}{2} \\ 4 \end{pmatrix} \tag{18}$$

- $\tau = 1$

$$\begin{array}{ccc} f(0,4) - (-\frac{9}{4}) & & -\frac{1}{2}(0 + \frac{9}{4}) \\ 16 + \frac{9}{4} & > & -\frac{9}{8}; \end{array}$$

- $\tau = \frac{1}{2}$

$$\begin{array}{ccc} f(0.75,2) - (-\frac{9}{4}) & & -\frac{1}{4} \cdot \frac{9}{4} \\ 0.8125 + \frac{9}{4} & > & -\frac{9}{16}; \end{array}$$

- $\tau = \frac{1}{4}$

$$\begin{array}{ccc} f(1.125,1) - (-\frac{9}{4}) & & -\frac{1}{8} \cdot \frac{9}{4} \\ -2.234 + \frac{9}{4} = 0.0016 & > & -\frac{9}{32}; \end{array}$$

- $\tau = \frac{1}{8}$

$$\begin{array}{ccc} f(1.3125,0.5) - (-\frac{9}{4}) & & -\frac{1}{16} \cdot \frac{9}{4} \\ -2.621 + \frac{9}{4} = -0.371 & \leq & -\frac{9}{64} = -0.14; \end{array}$$

Exercise 2 *Solve the following problem by using* **SUMT** *and taking* $x^{(0)} = (x_1^{(0)}, x_2^{(0)}) = (1, 1)$ *as a starting point*

$$\begin{aligned} \max_x f(x) &= x_1 x_2 \\ s.t. \quad & x_1^2 + x_2 - 3 \leq 0 \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned} \tag{19}$$

The solution is $x^* = (x_1^*, x_2^*) = (1, 2)$.

Solution : The starting point $x^{(0)} = (1, 1)$ is feasible and it is not on the boundary of the feasible region. We consider the following barrier function:

$$B(x) = \sum_{i=1}^{m} \frac{1}{b_i - g_i(x)} + \sum_{j=1}^{n} \frac{1}{x_j} \tag{20}$$

where $g_i(x), i = 1, \ldots, m$ are the constraint functions of the problem. In problem (19) there is only one ($m = 1$) constraint and two ($n = 2$) variables. The barrier function becomes:

$$B(x) = \frac{1}{3 - x_1^2 - x_2} + \frac{1}{x_1} + \frac{1}{x_2}. \tag{21}$$

The unconstrained problem which we use to approximate the initial problem (19) is defined as follows:

$$\max_x f(x) - \eta B(x) = x_1 x_2 - \eta\left(\frac{1}{3 - x_1^2 - x_2} + \frac{1}{x_1} + \frac{1}{x_2}\right). \tag{22}$$

The unconstrained problem (22) is solved for a decreasing sequence of values of the parameter η. That is in the beginning the parameter η is fixed to a specific value, say $\eta = a$, and problem (22) is solved. Next, the value of η is fixed to a smaller value, say $\eta = b < a$ and then again the corresponding problem (22) is solved. Since problem (22) is unconstrained it is much easier solved than the initial problem (19).

- Fix the parameter $\eta = 1$. Applying the steepest ascent method, starting from the point $x^{(0)} = (1, 1)$ we can find the maximum of the unconstrained problem (22), that is

  $$\max_x \quad x_1 x_2 - 1 \cdot \left(\frac{1}{3 - x_1^2 - x_2} + \frac{1}{x_1} + \frac{1}{x_2}\right). \tag{23}$$

  The solution of the above problem is $x^{(1)} = (0.90, 1.36)$.

- $x^{(1)}$ is an estimation of a local maximum of the initial constrained maximization problem (19). Also $x^{(1)}$ is a better estimation than the starting point $x^{(0)}$, as $f(x^{(1)}) > f(x^{(0)})$.

- Reducing the value of the parameter η by setting $\eta_1 = \eta\theta$ with $\theta = 0.01$ we have a new value of $\eta_1 = 0.01$.

- The parameter $\eta_1 = 0.01$. Applying the steepest ascent method, starting from the point $x^{(1)} = (0.90, 1.36)$ we can find the maximum of the unconstrained problem (22), that is

$$\max_x \quad x_1x_2 - 0.01 \cdot (\frac{1}{3 - x_1^2 - x_2} + \frac{1}{x_1} + \frac{1}{x_2}). \tag{24}$$

  The solution of the above problem is $x^{(2)} = (0.983, 1.933)$.

- $x^{(2)}$ is an estimation of a local maximum of the initial constrained maximization problem (19). Also $x^{(2)}$ is a better estimation than the previous point $x^{(1)}$, as $f(x^{(2)}) > f(x^{(1)})$.

- Reducing the value of the parameter η_1 by setting $\eta_2 = \eta_1\theta$ with $\theta = 0.01$ we have a new value of $\eta_2 = 0.0001$.

- The parameter $\eta_2 = 0.0001$. Applying the steepest ascent method, starting from the point $x^{(2)} = (0.983, 1.933)$ we can find the maximum of the unconstrained problem (22), that is

$$\max_x \quad x_1x_2 - 0.0001 \cdot (\frac{1}{3 - x_1^2 - x_2} + \frac{1}{x_1} + \frac{1}{x_2}). \tag{25}$$

  The solution of the above problem is $x^{(3)} = (0.998, 1.994)$.

- $x^{(3)}$ is an estimation of a local maximum of the initial constrained maximization problem (19). Also $x^{(3)}$ is a better estimation than the previous point $x^{(2)}$, as $f(x^{(3)}) > f(x^{(2)})$.

- By continuing this process (i.e. reducing the value of the parameter η and solving the corresponding unconstrained problem) we generate a sequence of points $x^{(k)}$, which converge to the local maximum of the initial constrained problem (19), which is $x^* = (1, 2)$.

SUMT is an iterative method which tries to find the optimum of constrained problem by solving a sequence of easier unconstrained problems. It is considered a very powerful technique and it is widely used for solving real life large scale problems. All the points it generates (approximations to the optimal solution) lie in the interior of the feasible region of the initial constrained problem. That is why it is often called the Interior Point Method. **SUMT** was invented by Fiacco and McCormick in 1968 and since 1984 there has been great interest in applying Interior Point Methods in linear and more recently in nonlinear optimization.