Problem 11.1: Interior Point Method for Linear Optimization

Consider the following (primal) Linear Optimization problem in standard form

$$(P): \ \min_{x} \ c^\top x \tag{1}$$
$$\text{subject to: } Ax = b \tag{2}$$
$$x \geq 0 \tag{3}$$

with variables $x \in \mathbb{R}^n$ and problem data $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, and $b \in \mathbb{R}^m$. The dual of the problem (1) – (3) can be written as

$$(D): \ \max_{v,u} \ b^\top v \tag{4}$$
$$\text{subject to: } A^\top v + u = c \tag{5}$$
$$u \geq 0 \tag{6}$$

with the dual variables $v \in \mathbb{R}^m$ and $u \in \mathbb{R}^n$. In this exercise, we want to solve the problem (1) – (3) using a primal-dual path following Interior Point Method (IPM). An optimal solution to both primal and dual problems satisfies the KKT conditions of P:

$$\begin{aligned} Ax &= b, & x &\geq 0, & &\text{(primal feasibility)} \\ A^\top v + u &= c, & u &\geq 0, & &\text{(dual feasibility)} \\ u^\top x &= 0. & & & &\text{(complementarity)} \end{aligned}$$

To solve (1) – (3) with the IPM, we will use a logarithmic barrier function. Adding this barrier function to the problem P, it becomes

$$(BP): \ \min. \ c^\top x - \mu \sum_{i=1}^{n} \log(x_i) \tag{7}$$
$$\text{subject to: } Ax = b, \tag{8}$$

where $\mu > 0$ is a suitable penalty parameter. The basic idea of IPMs is to initially solve BP with a large value of μ, iteratively reduce its value, and re-solve it until we are close enough to the optimum. We will solve (7) – (8) with a primal-dual path following variant of IPM in which we solve a single Newton step for each value of μ.

(a) Write the KKT conditions of the problem (7) – (8) and motivate a suitable stopping criterion for this IPM variant.

(b) Let $w = [x, v, u]^\top$, and let $H(\overline{w}) = 0$ denote the KKT system of part (a) for any solution $\overline{w} = [\overline{x}, \overline{v}, \overline{u}]^\top$ and penalty parameter μ. Let $J(\overline{w})$ be the Jacobian of $H(w)$ at $\overline{w}$, and let $d_w = (w - \overline{w}) = [d_x, d_v, d_u]^\top$ be the direction vector. Using this notation, apply Newton-Raphson method to solve $H(w) = 0$ at $\overline{w}$ and derive formulas for d_v, d_u, and d_x.

(c) At iteration k, we will solve $d_{w^{k+1}} = [d_{x^k}, d_{v^k}, d_{u^k}]^\top$ based on the update formulas derived in part (b) (i.e., Newton step), update the solution $w^{k+1} = w^k + d_{w^{k+1}}$, and set $\mu^{k+1} = \beta\mu^k$ for some $\beta \in (0, 1)$ until the stopping criterion derived in part (a) is satisfied.

To start the IPM, we need an initial primal solution x^0 (which does not necessarily need to be feasible). Propose a method to find a strictly feasible primal solution x^0 for (1) – (3).

(d) Solve the problem starting form the initial solution x^0 computed in part (c). Implementation of this IPM variant can be downloaded here.

Solution.

(a) The KKT conditions of the barrier problem (7) – (8) are

$$A x = b \qquad \text{(primal feasibility)} \tag{9}$$

$$A^\top v = c - \mu \left[\tfrac{1}{x_1}, \ldots, \tfrac{1}{x_n}\right]^\top \qquad \text{(dual feasibility)} \tag{10}$$

Since $\mu > 0$ and $x \geq 0$, we can interpret

$$u = \mu \left[\tfrac{1}{x_1}, \ldots, \tfrac{1}{x_n}\right] \geq 0$$

as estimates for Lagrangian dual variables u. Let $X \in \mathbb{R}^{n \times n}$ and $U \in \mathbb{R}^{n \times n}$ be defined as

$$X = \mathbf{diag}(x) = \begin{bmatrix} \ddots & & \\ & x_i & \\ & & \ddots \end{bmatrix} \quad \text{and} \quad U = \mathbf{diag}(u) = \begin{bmatrix} \ddots & & \\ & u_i & \\ & & \ddots \end{bmatrix}$$

and let $e = [1, \ldots, 1]^\top \in \mathbb{R}^n$ be a vector of ones. Using this notation, we can first define

$$u = \mu X^{-1} e \;\Rightarrow\; XUe = \mu e$$

and then rewrite the KKT conditions (9) – (10) as

$$A x = b \qquad \text{(primal feasibility)} \tag{11}$$

$$A^\top v + u = c \qquad \text{(dual feasibility)} \tag{12}$$

$$XUe = \mu e \qquad \text{(complementarity)} \tag{13}$$

To motivate a suitable stopping criterion, first notice that

$$XUe = \mu e \quad \Rightarrow \quad u_i x_i = \mu, \quad \text{for all } i = 1, \ldots, n.$$

This means that we have μ instead of 0 at the right hand side of each complementarity constraint. By multiplying both sides of (12) with $x^\top$ from left, we get

$$\begin{aligned} x^\top A^\top v + x^\top u &= x^\top c \qquad \text{or} \\ (Ax)^\top v + u^\top x &= x^\top c \qquad \text{or} \\ c^\top x - b^\top v &= u^\top x \end{aligned}$$

where $c^\top x - b^\top v = u^\top x$ is the duality gap for the current value of μ. Since

$$u^\top x = \sum_{i=1}^{n} u_i x_i = n\mu$$

we can use the value of $n\mu$ as a stopping condition: we can stop the Newton iterations if $n\mu < \epsilon$ for some predefined $\epsilon > 0$.

(b) Using the given notation, we want to solve

$$H(w) = \begin{bmatrix} Ax - b \\ A^\top v + u - c \\ XUe - \mu e \end{bmatrix} = 0$$

using the Newton-Raphson method. The Jacobian of $H(w)$ is

$$J(w) = \begin{bmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ U & 0 & X \end{bmatrix}$$

The Newton-Raphson method corresponds to solving the first order approximation at $\overline{w}$:

$$H(\overline{w}) + J(\overline{w})\underbrace{(w - \overline{w})}_{d_w} = 0$$

from which we get

$$J(\overline{w})d_w = -H(\overline{w})$$

or in matrix form

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ \overline{U} & 0 & \overline{X} \end{bmatrix} \begin{bmatrix} d_x \\ d_v \\ d_u \end{bmatrix} = - \begin{bmatrix} A\overline{x} - b \\ A^\top \overline{v} + \overline{u} - c \\ \overline{X}\,\overline{U}e - \mu e \end{bmatrix} \tag{14}$$

Notice that $A\overline{x} - b = 0$ and $A^\top \overline{v} + \overline{u} - c = 0$ do not necessarily hold until we are at optimum, since we make no assumptions about primal and dual feasibilities in this IPM variant.

Let us define primal and dual residual as

$$r_p = A\overline{x} - b \quad \text{and} \quad r_d = A^\top \overline{v} + \overline{u} - c, \tag{15}$$

respectively. Now, from (14) we get the following system of linear equations

$$Ad_x = -r_p \tag{16}$$
$$A^\top d_v + d_u = -r_d \tag{17}$$
$$\overline{U}d_x + \overline{X}d_u = \mu e - \overline{X}\,\overline{U}e \tag{18}$$

First, from (18) we get

$$d_u = \overline{X}^{-1}(\mu e - \overline{X}\,\overline{U}e) - \overline{X}^{-1}\overline{U}d_x \tag{19}$$

Substituting (19) to (17) yields

$$A^\top d_v + \overline{X}^{-1}(\mu e - \overline{X}\,\overline{U}e) - \overline{X}^{-1}\overline{U}d_x = -r_d$$

By separating the terms above, we have

$$\overline{X}^{-1}\,\overline{U}d_x = A^\top d_v + \overline{X}^{-1}(\mu e - \overline{X}\,\overline{U}e) + r_d$$

from which and we get a formula for d_x:

$$\begin{aligned} d_x &= (\overline{X}^{-1}\,\overline{U})^{-1}A^\top d_v + (\overline{X}^{-1}\,\overline{U})^{-1}\overline{X}^{-1}(\mu e - \overline{X}\,\overline{U}e) + (\overline{X}^{-1}\,\overline{U})^{-1}r_d \\ &= \overline{U}^{-1}\overline{X}A^\top d_v + \overline{U}^{-1}\overline{X}\,\overline{X}^{-1}(\mu e - \overline{X}\,\overline{U}e) + \overline{U}^{-1}\overline{X}r_d \\ &= \overline{U}^{-1}\overline{X}A^\top d_v + \overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e + \overline{X}r_d) \end{aligned} \tag{20}$$

Since $Ad_x = -r_p$ by (16), we can solve d_v from (20) by first writing

$$Ad_x = A\overline{U}^{-1}\overline{X}A^\top d_v + A\overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e + \overline{X}r_d) \tag{21}$$
$$= -r_p \tag{22}$$

Based on (21) – (22), we get the expression

$$A\overline{U}^{-1}\overline{X}A^\top d_v = -A\overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e + \overline{X}r_d) - r_p$$

from which we can solve d_v as

$$d_v = -(A\overline{U}^{-1}\overline{X}A^\top)^{-1}[A\overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e + \overline{X}r_d) + r_p] \tag{23}$$

After computing d_v from (23), we can solve d_u by simply substituting d_v to (17):

$$d_u = -A^\top d_v - r_d \tag{24}$$

Finally, after computing d_u from (24), we can solve d_x by substituting d_u to (18):

$$d_x = \overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e - \overline{X}d_u) \tag{25}$$

Thus, we get the following equations for d_v, d_u, and d_x that are updated in sequence at each Newton iteration k to get the direction vectors for iteration $k+1$:

$$\begin{aligned}
d_v &= -(A\overline{U}^{-1}\overline{X}A^\top)^{-1}[A\overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e + \overline{X}r_d) + r_p] \\
d_u &= -A^\top d_v - r_d \\
d_x &= \ \overline{U}^{-1}(\mu e - \overline{X}\,\overline{U}e - \overline{X}d_u)
\end{aligned}$$

(c) We need a strictly feasible initial primal solution x^0 that satisfies (2) – (3). The solution x^0 must be strictly feasible because we have the sum of logarithms in the objective (7). One way to obtain such a solution is to solve the following linear optimization problem:

$$\begin{aligned}
\text{min. } & t && (26)\\
\text{subject to: } & Ax = b && (27)\\
& t \geq 1 - x_i, \quad \forall i = 1, \dots, n && (28)\\
& t \geq 0 && (29)
\end{aligned}$$

with variables $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$. The problem (1) – (3) is strictly feasible if and only if the problem (26) – (29) has an optimal solution (t^*, x^*) with $t^* < 1$. Moreover, $x^0 = x^*$ is then a strictly feasible primal solution.

(d) The IPM algorithm is implemented in this Jupyter notebook.