SF2812 Applied linear optimization, final exam
March 14 2024 08.00–13.00
Brief solutions

1. (a) To keep the notation as simple as possible, I will use the notation x_1, x_2, x_3, x_4 for the variables representing the amount produced of each product. Similarly, I use xe_1, xe_2, xe_3 for the excess variables. The problem can then be written as

$$
\begin{aligned}
\text{minimize} \quad & -8x_1 - 7x_2 - 8x_3 - 20x_4 \\
\text{subject to} \quad & 2x_1 + 4x_2 + 5x_3 + 8x_4 + xe_1 = 15, \\
& 4x_1 + 1x_2 + 7x_3 + 4x_4 + xe_2 = 10, \\
& 9x_1 + 7x_2 + 8x_3 + 15x_4 + xe_3 = 80, \\
& x_1, x_2, x_3, x_4 \geq 0, \\
& xe_1, xe_2, xe_3 \geq 0.
\end{aligned} \tag{1}
$$

(b) From the GAMS output we get

$$
x = \begin{pmatrix} 0.8333 \\ 0 \\ 0 \\ 1.6667 \end{pmatrix}, \quad xe = \begin{pmatrix} 0 \\ 0 \\ 47.50 \end{pmatrix}, \quad y = \begin{pmatrix} -2 \\ -1 \\ 0 \end{pmatrix}, \quad s = \begin{pmatrix} s_x \\ s_{xe} \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 9 \\ 0 \\ 2 \\ 1 \\ 0 \end{pmatrix}.
$$

(c) From the GAMS output, we can see that both raw materials 1 and 2 are limiting (all is used). The dual variables y_1 and y_2 show that increasing the amount of raw materials 1 and 2 would improve the solution. But, the dual variable y_1 is more negative indicating that increasing raw material 1 would be more beneficial.

(d) Product 2, as it has a smaller reduced cost (s_2).

(e) We know that the reduced costs $s_N = c_N - N^T y$. The question is really, how much smaller can we make c_n before s_n becomes negative?

For product 2 we have $s_{x_2} = 2$ and $c_2 = -7$. If we make the objective coefficient for product 2 more than two units smaller then the reduced cost becomes negative. Thus if the coefficient for product 2 is smaller than -9, the solution will change.

Similarly, we get that the coefficient for product 3 needs to be smaller than -17 for the solution to change.

2. (a) For the solution to be optimal it needs to be both primal and dual feasible and satisfy $x^T s = 0$. For the solution provided by the friend, we get $x^T s = 2$. Therefore, it cannot be optimal!

(b) With x_1, x_2, x_5 as basic, x_3, x_4 as nonbasic variables we get

$$
B = \begin{pmatrix} -4 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.
$$

We can easily calculate the calues for the basic variables as

$$x_B = B^{-1}b = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}.$$

Next, we calculate y from $B^T y = c_B$ and get

$$y = \begin{pmatrix} 0.25 \\ -1.25 \\ 0 \end{pmatrix}.$$

We then calculate the reduced costs $s_N = c_N - N^T y$, and get

$$s_N = \begin{pmatrix} -0.25 \\ 1.25 \end{pmatrix}.$$

We should, thus move x_3 into the basis. We next calculate the search direction p_B for the basic variables by

$$p_B = -B^{-1}N_1 = \begin{pmatrix} 0.25 \\ 0 \\ -0.5 \end{pmatrix}. \qquad \text{(where } N_1 \text{ is the first column of } N\text{)}$$

For the nonbasic variables the search direction p_N is simply given by

$$p_N = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

When moving in this direction x_5 is the first variable to reach zero (leaves the basis), and we get the maximal step length $\alpha_{max} = 2$. We can then update the variables by

$$x = \begin{pmatrix} 1 \\ 4 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} 0.25 \\ 0 \\ 1 \\ 0 \\ -0.5 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 4 \\ 2 \\ 0 \\ 0 \end{pmatrix}.$$

Now we still need to check if this is optimal. With x_1, x_2, x_3 as basic variables we get

$$B = \begin{pmatrix} -4 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 1 & 0 \end{pmatrix}.$$

We calculate y as before and get

$$y = \begin{pmatrix} 0 \\ -0.5 \\ -0.5 \end{pmatrix},$$

and we get the reduced costs

$$s_N = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

As the reduced costs are positive the solution is optimal.

(c) From the simplex calculations we got all that we need (but, we need to remember that $s_B = 0$). The dual variables are thus

$$y = \begin{pmatrix} 0 \\ -0.5 \\ -0.5 \end{pmatrix}, \quad s = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.5 \\ 0.5 \end{pmatrix}.$$

3. See material from lecture 10, and chapter 7.4 in the book.

An ideal solution should start from the master problem. Then explain how the reduced costs for the master problem can be determined and how the smallest reduced costs can be found by solving an LP problem that gives you the smallest reduced costs and the new extreme point to enter the basis.

4. (a) We are comparing the inequality

$$x_3 + 2x_4 + 2x_5 \leq 3, \tag{2}$$

with

$$x_3 + x_4 + x_5 \leq 2, \tag{3}$$
$$x_4 + x_5 \leq 1. \tag{4}$$

As all the variables x_3, x_4, and x_5 are binary, it is clear that the feasible integer-valued solutions to (2) are the same as the feasible integer-valued solutions to (3) and (4). But, when we relax integrality and allow for fractional solution then the feasible sets will not be the same as we will show.

First, note that we can obtain (2) by simply adding (3) and (4) together. Therefore, any solution that satisfies (3) and (4) must automatically satisfy (2). It is now clear that (3) and (4) cannot give a weaker continuous relaxation than (2). But, we still need to show that it is stronger.

Consider the fractional point $x_3 = 0, x_4 = 0.75, x_5 = 0.75$. This point satisfies (2) but clearly violates (4). Thus we have shown that (3) and (4) give a smaller set of feasible fractional solutions than (2).

(b) There are several ways to formulate this as a MILP and there are several ways of deriving the constraints needed. Here, I will show one approach. You don't need to follow these steps to get a correct solution. But, I think this is the most systematic and clear way of deriving the constraints. In the exam, I didn't expect all these steps, but I included it here to make the explanation as clear as possible.

First, we focus on node n1. For n1 it is clear that $z_1 \geq 0$. We can also include the constraint $z_1 \geq -2x_1 + x_2 + 3$.
We can consider the node as the disjunctive constraint

$$\begin{bmatrix} z_1 = 0 \end{bmatrix} \vee \begin{bmatrix} z_1 = -2x_1 + x_2 + 3 \\ z_1 \geq 0 \end{bmatrix}. \tag{5}$$

The disjunctive constraint simply says that either $z_1 = 0$ or $(z_1 = -2x_1 + x_2 + 3, \quad z_1 \geq 0)$. We will rewrite the equality constraints as two inequality constraints (we know how to deal with inequality constraints with the big-M approach). We then get

$$\begin{bmatrix} z_1 \leq 0 \\ z_1 \geq 0 \end{bmatrix} \vee \begin{bmatrix} z_1 \leq -2x_1 + x_2 + 3 \\ z_1 \geq -2x_1 + x_2 + 3 \\ z_1 \geq 0 \end{bmatrix}. \tag{6}$$

We can directly apply the big-M approach to the disjunctive constraint above to get a mixed-integer formulation of the node. But, I want to clean it up a bit to avoid unnecessary constraints. (But this is not necessary, you will still get a correct solution without this step)
As first mentioned, we know that we can include the constraints $z_1 \geq 0$ and $z_1 \geq -2x_1 + x_2 + 3$. So all the constraints we have are

$$\begin{bmatrix} z_1 \leq 0 \\ z_1 \geq 0 \end{bmatrix} \vee \begin{bmatrix} z_1 \leq -2x_1 + x_2 + 3 \\ z_1 \geq -2x_1 + x_2 + 3 \\ z_1 \geq 0 \end{bmatrix}, \tag{7}$$

$$z_1 \geq -2x_1 + x_2 + 3, \tag{8}$$

$$z_1 \geq 0. \tag{9}$$

Note that we have constraints (8) and (9) both as separate constraints and within the disjunctive constraint (7). We can remove these from the disjunctive constraint to reduce the number of constraints and we are then left with

$$\begin{aligned} &\begin{bmatrix} z_1 \leq 0 \end{bmatrix} \vee \begin{bmatrix} z_1 \leq -2x_1 + x_2 + 3 \end{bmatrix}, \\ &z_1 \geq -2x_1 + x_2 + 3, \\ &z_1 \geq 0. \end{aligned}$$

By introducing two binary variables b_0 and b_1 we can use the big-M formulation to represent the disjunctive constraint, and we get

$$\begin{aligned} &z_1 \leq 0 + M_1(1 - b_0), \\ &z_1 \leq -2x_1 + x_2 + 3 + M_2(1 - b_1), \\ &z_1 \geq -2x_1 + x_2 + 3, \\ &z_1 \geq 0, \\ &b_0 + b_1 = 1. \end{aligned}$$

As a final step, we need to determine suitable values for $M1$ and $M2$, and to do so we need to analyze the expression $-2x_1 + x_2 + 3$. Since both x_1 and x_2 are in restricted to be within the interval $[-5, 5]$ we can easily determine the following upper and lower bounds on the expression

$$-12 \leq -2x_1 + x_2 + 3 \leq 18. \tag{10}$$

Thus, we can use $M_1 = 18$ and $M_2 = 12$. The MILP representation of node n1 is then given by

$$z_1 \leq 18(1 - b_0), \tag{11}$$
$$z_1 \leq -2x_1 + x_2 + 3 + 12(1 - b_1), \tag{12}$$
$$z_1 \geq -2x_1 + x_2 + 3, \tag{13}$$
$$z_1 \geq 0, \tag{14}$$
$$b_0 + b_1 = 1. \tag{15}$$

If we want to, we can also eliminate one of the binary variables using the equation $b_0+b_1 = 1$. For example, if we solve the equation for b_0 and substitute in $b_0 = 1 - b_1$ then we get

$$z_1 \leq 18b_1, \tag{16}$$
$$z_1 \leq -2x_1 + x_2 + 3 + 12(1 - b_1), \tag{17}$$
$$z_1 \geq -2x_1 + x_2 + 3, \tag{18}$$
$$z_1 \geq 0. \tag{19}$$

But, either formulation is ok! I will use the latter formulation, simply because there are fewer things to write down.

Using a similar approach we can get the following bound for the expression in n2

$$-21 \leq 3x_1 - x_2 - 1 \leq 19. \tag{20}$$

Then we know that $0 \leq z_1 \leq 18$ and $0 \leq z_2 \leq 19$, and we can determine bounds for the expression in n3 according to

$$-23 \leq 2z_1 - z_2 - 4 \leq 32. \tag{21}$$

I will use the formulation with just one binary variable, per node and I will denote the binary variable for node n2 by b_2 and the one for n3 by b_3. Problem

(P1) can then be written as

$$
\begin{aligned}
\text{minimize} \quad & y \\
\text{subject to} \quad & Ax \leq b, \\
& z_1 \leq 18b_1, \\
& z_1 \leq -2x_1 + x_2 + 3 + 12(1 - b_1), \\
& z_1 \geq -2x_1 + x_2 + 3, \\
& z_2 \leq 19b_2, \\
& z_2 \leq 3x_1 - x_2 - 1 + 21(1 - b_2), \\
& z_2 \geq 3x_1 - x_2 - 1, \\
& y \leq 32b_3, \\
& y \leq 2z_1 - z_2 - 4 + 23(1 - b_3), \\
& y \geq 2z_1 - z_2 - 4, \\
& -5 \leq x_i \leq 5, \quad i = 1, 2, \\
& z_1, z_2, y \in I\!R_+, \\
& b_1, b_2, b_3 \in \{0, 1\}, \\
& x \in I\!R^n, \ x_i \in \mathbb{Z} \quad \forall i \in \mathcal{I}.
\end{aligned}
\tag{22}
$$

Here, I assume the constraints from question a) are all in the constraints $Ax \leq b$. This is written with as few constraints and variables as possible, and it is ok if you have more variables and constraints.

5. (a) i. First remember $||u||_1 = \sum_{i=1}^{n} |u_i|$. By introducing two new positive variables u_i^+ and u_i^- for each u_i we can write the constraint with the 1-norm as

$$\sum_{i=1}^{n} u_i^+ + u_i^- \leq \epsilon, \tag{23}$$

$$u = u^+ - u^-, \tag{24}$$

$$u^+ \geq 0, \quad u^- \geq 0. \tag{25}$$

The new variables u^+ and u^- are here used to represent the absolute values.

Alternatively one could introduce just one new variable $\bar{u}_i$ for each u_i along with the constraints $u_i \leq \bar{u}_i$ and $-u_i \leq \bar{u}_i$. The norm constraint can then be written as $\sum_{i=1}^{n} \bar{u}_i \leq \epsilon$. This will result in a slightly more complicated (but correct dual) and, therefore, I will go with the first option. However, you can solve the problem without any issues with both approaches.

The constraint problems can thus be written as

$$
\begin{array}{ll}
\text{minimize} & a_i^T(x-u) \\
\text{subject to} & \sum_{i=1}^{n} u_i^+ + u_i^- \leq \epsilon, \\
& u = u^+ - u^-, \\
& u^+ \geq 0, \quad u^- \geq 0, \\
& u^+, u^-, u \in \mathbb{R}^n.
\end{array}
\tag{26}
$$

This is one correct solution to the first question.

ii. First, we will slightly rewrite the constraint problems (26) to simplify the derivation of the dual (but, these steps are not needed). By introducing a vector of n ones $e = [1 \;\; 1 \;\; \dots \;\; 1]^T$, we can write the first constraint in (26) as $e^T(u_i^+ + u_i^-) \leq \epsilon$,. Furthermore, by introducing a slack variable s_1 we can write the inequality constraint as an equality constraint. (The reason behind writing the constraint as an equality constraint; we mainly looked at how to derive the dual for equality or greater than constraints during the lectures). We will thus consider the constraint problems (26) written in the form

$$\begin{aligned} \text{minimize} \quad & a_i^T(x-u) \\ \text{subject to} \quad & e^T(u_i^+ + u_i^-) + s_1 = \epsilon, \\ & u = u^+ - u^-, \\ & u^+ \geq 0, \quad u^- \geq 0, \quad s_1 \geq 0 \\ & u^+, u^-, u \in I\!R^n, \quad s_1 \in I\!R. \end{aligned} \tag{27}$$

Next, we need to derive the dual of the problem above. We introduce the dual variables λ_i for the constraint $u = u^+ - u^-$ and the dual variable μ_i for the constraint $e^T(u_i^+ + u_i^-) + s_1 = \epsilon$. We can then define the Lagrangian dual function as

$$\varphi(\lambda, \mu) = \underset{u \in I\!R^n, u^+ \in I\!R^n_+, u^- \in I\!R^n_+, s_1 \in I\!R_+}{\text{minimize}} \; a_i^T(x-u) - \lambda^T(u - u^+ + u^-) - \mu\left(e^T(u_i^+ + u_i^-) + s_1 - \epsilon\right).$$

(Remember, the dual function can be defined in different ways and you will still get a correct dual problem.)

By rearranging the terms in the dual function we can write it as

$$\varphi(\lambda, \mu) = \underset{u \in I\!R^n, u^+ \in I\!R^n_+, u^- \in I\!R^n_+, s_1 \in I\!R_+}{\text{minimize}} \; a_i^T x + \mu\epsilon - (a_i + \lambda)^T u + (\lambda - \mu e)^T u^+ - (\lambda + \mu e)^T u^- - \mu s_1.$$

We then get

$$\varphi(\lambda, \mu) = \begin{cases} a_i^T x + \mu\epsilon & \text{if } a_i + \lambda = 0, \; \lambda \geq \mu e, \; -\lambda \geq \mu e, \; \mu \leq 0 \\ \infty & \text{otherwise} \end{cases}$$

and we can write the dual problem as

$$\begin{aligned} \text{maximize} \quad & a_i^T x + \mu\epsilon \\ \text{subject to} \quad & a_i + \lambda = 0, \\ & \lambda \geq \mu e, \\ & -\lambda \geq \mu e, \\ & \mu \leq 0. \end{aligned} \tag{28}$$

(b) Note that we have one separate constraint problem and a corresponding dual problem for each of the m constraints. But, these problems and their duals all have the same form. The dual of the constraint problems are

$$
\begin{aligned}
\text{maximize} \quad & a_i^T x + \mu_i \epsilon \\
\text{subject to} \quad & a_i + \lambda = 0, \\
& \lambda_i \ge \mu_i e, \\
& -\lambda_i \ge \mu_i e, \\
& \mu_i \le 0, \\
& \lambda_i \in I\!R^n, \ \mu_i \in I\!R.
\end{aligned} \tag{29}
$$

Now, note that each of these dual problems has different sets of variables. Since each problem is feasible and strong duality holds, we know that we can replace the optimization problems in the constraints with their dual problems (they obtain the same optimal objective value). We can then change the constraints problems from minimization problems into maximization problems. Problem $(R - LP)$ can then be written as

$$
\begin{aligned}
\underset{x \in I\!R^n}{\text{minimize}} \quad & c^T x \\
\text{subject to} \quad & \begin{pmatrix} \text{maximize} \quad a_i^T x + \mu_i \epsilon \\ \text{subject to} \quad \mathrm{a_i} + \lambda = 0, \\ \lambda_i \ge \mu_i e, \\ -\lambda_i \ge \mu_i e, \\ \mu_i \le 0, \\ \lambda_i \in I\!R^n, \ \mu_i \in I\!R \end{pmatrix} \ge b_i, \quad i = 1, \dots m.
\end{aligned}
$$

Note that each constraint problem has different sets of variables. From the duality theory, we know that any feasible solution to the dual constraint problems (29) provides a lower bound of the optimal objective value of the corresponding constraint problem (26). More specifically, for any λ_i, μ_i satisfying the constraints in (29) we get

$$
\begin{aligned}
a_i^T x + \mu_i \epsilon \le \quad \text{minimize} \quad & a_i^T (x - u) \\
\text{subject to} \quad & \sum_{i=1}^{n} u_i^+ + u_i^- \le \epsilon, \\
& u = u^+ - u^-, \\
& u^+ \ge 0, \quad u^- \ge 0, \\
& u^+, u^-, u \in I\!R^n.
\end{aligned}
$$

Dropping the maximization operator from the constraint problems could, therefore, only result in a potentially stronger constraint (but all constraints from the dual problem are needed to ensure that $a_i^T x + \mu_i \epsilon$ gives a lower bound). Since strong duality holds we know that $a_i^T x + \mu_i \epsilon$ will be equal to the minimum for the optimal dual variables. Now, we can observe that the max operator in the constraints is redundant. If $a_i^T x + \mu_i \epsilon$ is less than its maximum, it means we are satisfying the constraint “more than needed”.

As long as we are taking the constraints of each constraint problem into consideration we are allowing $a_i^T x + \mu_i \epsilon$ to obtain its maximum or a smaller value (and a smaller value just results in a stricter constraint). When we skip the maximization operator in the constraints we are left with

$$
(R-LP^*) \qquad
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & a_i^T x + \mu_i \epsilon \geq b_i, \quad i = 1, \dots m. \\
& a_i + \lambda = 0, \quad i = 1, \dots m. \\
& \lambda_i \geq \mu_i e, \quad i = 1, \dots m. \\
& -\lambda_i \geq \mu_i e, \quad i = 1, \dots m. \\
& \mu_i \leq 0, \quad i = 1, \dots m. \\
& \lambda_i \in I\!R^n, \ \mu_i \in I\!R, \quad i = 1, \dots m. \\
& x \in I\!R^n.
\end{array}
$$

This linear programming problem is equivalent to $(R-LP)$ in the sense they have the same minimizer x.