OPT Problem Sheet 5 Solutions

Duality and Sensitivity.

1. (a) `False.` $\boldsymbol{c}$ does not affect feasibility of $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq 0$.

   (b) `True.` The solution $\boldsymbol{y} = A_B^{-T}\boldsymbol{c}_B$ is still feasible for the dual, so the basis remains optimal for the primal, as long as it stays feasible.

   (c) `True.` By weak duality. The value of the dual is always $\geq$ the value of the primal, for any pair of feasible solutions $(\boldsymbol{x}, \boldsymbol{y})$ for the pair (Primal, Dual). So. if the values are equal, both must be optimal.

2. This is just a matter of terminology. Reduced costs are slacks in the dual constraints, corresponding to the slack variables introduced into the MP. These slacks equal the solution $\boldsymbol{y}$ of the dual. Namely,

   MP: $\texttt{max}\ \boldsymbol{c} \cdot \boldsymbol{x},\ A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{R}^n_+,\ \boldsymbol{b} \in \mathbb{R}^m$.

   Add slack variable $\boldsymbol{s} \in \mathbb{R}^m_+$. Let $\tilde{\boldsymbol{x}} = (\boldsymbol{x}, \boldsymbol{s})$, $\tilde{A} = [A\ I]$, $\tilde{\boldsymbol{c}} = (\boldsymbol{c}, \boldsymbol{0})$, where I is $m \times m$ identity matrix. Canonical form:

   $$\tilde{A}\tilde{\boldsymbol{x}} = \boldsymbol{b},\ \tilde{\boldsymbol{x}} \in \mathbb{R}^{n+m}_+,\ \texttt{max}\ \tilde{\boldsymbol{c}} \cdot \tilde{\boldsymbol{x}}.$$

   Dual: $\tilde{A}^T\boldsymbol{y} \geq \tilde{\boldsymbol{c}}$, $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$. The group of the dual inequalities, corresponding to the identity I in $\tilde{A} = [A\ I]$ simply reads $\boldsymbol{y} \geq \boldsymbol{0}$. This holds for any feasible solution $\boldsymbol{y}$ of the dual.

   Now, by definition, if $\boldsymbol{y}$ is the optimal solution for the dual, then the difference between the left and right-hand sides in the dual inequalities is called the reduced cost of the corresponding primal variable. On the other hand, $\boldsymbol{y}$ itself has the dimension of $\boldsymbol{b}$, i.e. the constraints' vector, and $\boldsymbol{y}$ is called the shadow price of the constraints. So the two notions – reduced cost of the slack variables, which is $\boldsymbol{y} - \boldsymbol{0}$ and the shadow price, which is $\boldsymbol{y}$ itself, coincide. In particular, if at the end of the day one of the slack variables is basic, i.e. nonzero, so the corresponding inequality is realised with a slack, then its shadow price equals zero (being the reduced cost of a basic variable).

   Note: You can relate this to the dual simplex method: if there is a $\leq$ constraint, its slack variable would invariably be basic for the initial tableau, as the corresponding column in the initial tableau will invariably be the column of the identity matrix. The shadow price if the constraint will be what one ends up having in the objective row in this column after the final tableau has been obtained.

   As for the $\geq$ constraints, the reduced cost of the corresponding excess variable is NOT generally equal to the constraint's shadow price (see e.g. problem 1(a) in the first group). Instead, one has to read out of the final tableau the objective row entry from the column, which in the original tableau was the corresponding column of the unit matrix (this column would often be represented by an artificial variable that one would keep track of, despite switching from Phase I to Phase II).

   In other words, the shadow price of the ith constraint in the dual simplex method is given by the final tableau objective row entry in the row that used to be the ith row of the identity matrix in the original tableau.

3. MP: $\texttt{max}\ \boldsymbol{c} \cdot \boldsymbol{x},\ A\boldsymbol{x} \leq \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{R}^n_+,\ \boldsymbol{b} \in \mathbb{R}^m$.

   Add slack variable $\boldsymbol{s} \in \mathbb{R}^m_+$. Let $\tilde{\boldsymbol{x}} = (\boldsymbol{x}, \boldsymbol{s})$, $\tilde{A} = [A\ I]$, $\tilde{\boldsymbol{c}} = (\boldsymbol{c}, 0)$, where I is $m \times m$ identity matrix. Canonical form:

   $$\tilde{A}\tilde{\boldsymbol{x}} = \boldsymbol{b},\ \tilde{\boldsymbol{x}} \in \mathbb{R}^{n+m}_+,\ \texttt{max}\ \tilde{\boldsymbol{c}} \cdot \tilde{\boldsymbol{x}}.$$

   Dual: $\tilde{A}^T\boldsymbol{y} \geq \tilde{\boldsymbol{c}}$, $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$. Using the above form of $\tilde{A}$ and $\tilde{\boldsymbol{c}}$, the dual feasibility boils down to $A^T\boldsymbol{y} \geq \boldsymbol{c},\ \boldsymbol{y} \geq 0$, with the same objective $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$.

   DP: $\texttt{min}\ \boldsymbol{c} \cdot \boldsymbol{x},\ A\boldsymbol{x} \geq \boldsymbol{b},\ \boldsymbol{x} \in \mathbb{R}^n_+,\ \boldsymbol{b} \in \mathbb{R}^m$.

   Add excess variable $\boldsymbol{e} \in \mathbb{R}^m_+$. Let $\tilde{\boldsymbol{x}} = (\boldsymbol{x}, \boldsymbol{e})$, $\tilde{A} = [A\ \ -I]$, $\tilde{\boldsymbol{c}} = (-\boldsymbol{c}, 0)$. Canonical form:

   $$\tilde{A}\tilde{\boldsymbol{x}} = \boldsymbol{b},\ \tilde{\boldsymbol{x}} \in \mathbb{R}^{n+m}_+,\ \texttt{max}\ \tilde{\boldsymbol{c}} \cdot \tilde{\boldsymbol{x}}.$$

Dual: $\tilde{A}^T \boldsymbol{y} \geq \tilde{\boldsymbol{c}}$, $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$. Using the above form of $\tilde{A}$ and $\tilde{\boldsymbol{c}}$, the dual feasibility boils down to $A^T \boldsymbol{y} \geq -\boldsymbol{c}$, $\boldsymbol{y} \leq 0$. So, change $\boldsymbol{y}$ to $-\boldsymbol{y}$, then the dual is $A^T \boldsymbol{y} \leq \boldsymbol{c}$, $\boldsymbol{y} \geq 0$, with the same objective $\texttt{max}\ \boldsymbol{b} \cdot \boldsymbol{y}$.

4. Write the dual $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$, $A^T \boldsymbol{y} \geq \boldsymbol{c}$, $\boldsymbol{y} \in \mathbb{R}^m$, $\boldsymbol{c} \in \mathbb{R}^n$ as canonical form by changing $\texttt{min}\ \boldsymbol{b} \cdot \boldsymbol{y}$ to $\texttt{max}\ -\boldsymbol{b} \cdot \boldsymbol{y}$, writing $\boldsymbol{y} = \boldsymbol{u} - \boldsymbol{v}$, and adding the excess variable $\boldsymbol{e}$:

$$\texttt{max}\ -\boldsymbol{b} \cdot \boldsymbol{u} + \boldsymbol{b} \cdot \boldsymbol{v} + \boldsymbol{0} \cdot \boldsymbol{e},\ A^T \boldsymbol{u} - A^T \boldsymbol{v} - I\boldsymbol{e} = \boldsymbol{c},\ \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^m_+,\ \boldsymbol{e} \in \mathbb{R}^n_+,$$

where I is $n \times n$ identity matrix. The dual to this canonical form is

$$\texttt{min}\ \boldsymbol{c} \cdot \boldsymbol{z},\ A\boldsymbol{z} \geq -\boldsymbol{b},\ -A\boldsymbol{z} \geq \boldsymbol{b},\ -\boldsymbol{z} \geq 0.$$

Let now $\boldsymbol{z} = -\boldsymbol{x}$ to get back the primal canonical form

$$\texttt{max}\ \boldsymbol{c} \cdot \boldsymbol{x},\ A\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq 0.$$

5. Introduce a slack variable x_6 and excess variable x_7, put the problem into canonical form, with the 3×7 matrix A. Check that the given solution implies $x_6 = x_7 = 0$.

Now consider the dual inequalities as *equations* for the basic components. The number of unknowns should be the number of constraints, the coefficients in the system of equations in each equation - simply the coefficients in the basic columns of A. The right-hand side are the basic components of $\boldsymbol{c} = (1, 2, 3, -2, 8, 0, 0)$. I.e. solve the system of equations

$$y_1 - y_2 + 3y_3 = 1, -y_1 + 3y_2 + y_3 = 3, 2y_1 + y_2 - 2y_3 = 8,$$

with the unknowns (y_1, y_2, y_3). This results in a solution $(y_1, y_2, y_3) = (3, 2, 0)$.

Now check optimality of $\boldsymbol{x} = (2.75, 0, 4.75, 0, 4.5, 0, 0)$, i.e. whether the solution found satisfies the dual inequalities, for the free components of A. These inequalities are:

$$3y_1 - y_2 + 2y_3 \geq 1,\ 2y_1 - y_3 \geq -2,\ y_2 \geq 0,\ y_3 \leq 0.$$

They are clearly satisfied by the above $(y_1, y_2, y_3) = (3, 2, 0)$.

Conclusion: the solution $\boldsymbol{x}$ of the primal is optimal, the shadow prices of the constraints are $(y_1, y_2, y_3) = (3, 2, 0)$, the reduced costs of the free variables (x_2, x_4) equal $(6, 6)$ (the margin in the corresponding free dual inequalities), and of (x_5, x_6) are $(2, 0)$. Note that the reduced cost of a slack variable is always non-negative and equal to the shadow price of the underlying constraint; for an excess variable its reduced cost is always non-positive and equal to the shadow price of the underlying constraint.

6. See solutions to set 4.

Geometry of LP

1. Interior: $\{(x, y) : x^2 + y^2 > 1\}$, boundary $\{(x, y) : x^2 + y^2 = 1\}$. There are at least two ways to prove that the set $X = \{(x, y) : x^2 + y^2 \geq 1\}$ is closed.

First way: prove that its complement $X^c = \{(x, y) : x^2 + y^2 < 1\}$ is open. I.e., for any $(x, y) : x^2 + y^2 < 1$ show that there is a small disk around (x, y) which is fully contained in X^c. This is certainly true: for any $(x, y) \in X^c$ let $\delta = 1 - (x^2 + y^2)$. Then $1 \geq \delta > 0$, and now take the disk around (x, y) with the radius $\delta/100$. Then this disk lies fully in X^c, as

$$(x \pm \delta/100)^2 + (y \pm \delta/100)^2 \leq 1 - \delta + \delta/25 + \delta^2/5000 < 1.$$

Second way – the limit point argument. Prove that any for any sequence (x_n, y_n), converging to (x, y), and such that $\forall n$, $x_n^2 + y_n^2 \geq 1$, the same is true for (x, y). This follows from general properties of limits. Let $a_n = x_n^2 + y_n^2$. Then $a = \lim a_n = x^2 + y^2$. And as every $a_n \geq 1$, then $a \geq 1$.

2. The set of all limit points is the closure of the set, i.e. $\{(x, y) : x^2 - y^2 \geq 1\}$.

3. Closure: the whole unit circle. Interior: empty – there is no open disk contained in the unit circle. Boundary=closure\interior=the whole unit circle.

   Note: to prove the fact that the closure is the whole unit circle is a bit tricky, and you need to use irrationality of 2π for that, as well as compactness of the circle. Indeed, let P_n be the point on the unit circle, the member of the sequence. We have to prove that for any point P on the unit circle, there is some P_n arbitrarily close to P. Suppose, this is not true. Then there is a P on the unit circle and some interval G around P, with no points of the sequence $\{P_n\}$ inside this interval. Call such an interval a gap. Now, of all gaps take the largest, call it G_0. If G_0 is a gap, then G_0, rotated by the angle 1 is also a gap. Call it G_1. Rotate again – get the gap G_2, and so on. If you do it many enough times, there are two options. The first one is that for some N, $G_0 = G_N$, i.e. after rotating the gap N times, it has come back straight onto itself. This is impossible, however, because it would mean that 2π is a rational number. So, the gap would never come back exactly onto itself. But then, for large enough N, the gaps G_0 and G_N will overlap, thus creating a bigger gap, whose size is strictly greater than the size of G_0, which is a contradiction with how G_0 has been chosen.

4. The closure of X is: (i) the set of all limit points of X, and (ii) the smallest (in the sense of set-theoretical inclusion) closed set, containing X.

   Let's first prove that there is no closed X' containing X which is smaller than the set of all limit points of X. Indeed, otherwise X' contains X, but does not contain some limit point P of X. The complement of X' is open, so X' also does not contain some open neighbourhood of P, of radius, say δ. But then P cannot be a limit point of X, because no point of X gets closer than the distance δ to P.

   Now let's show that the set of all limit points of X is actually closed then. Take the closure $\bar{X}$ of X as the smallest closed set containing X. We already know now that it should contain all the limit points of X. If it contains anything else, i.e. some point P which is *not* a limit point of X, then once again there is a little open ball $B(P,\delta)$ of radius δ around P, containing no points of X. Let $\bar{B}(P,\delta/2)$ be the closed ball centered at P, of radius $\delta/2$. Then $X' = \bar{X} \setminus \bar{B}(P,\delta/2)$ is a closed set, and we have proper inclusions $X \subset X' \subset \bar{X}$. This violates the premise that $\bar{X}$ is the smallest closed set containing X.

5. Take two points with radius vectors $\boldsymbol{r}_1 = (x_1, y_1, z_1)$ and $\boldsymbol{r}_2 = (x_2, y_2, z_2)$. A point between them has a radius vector $\boldsymbol{r} = \theta\boldsymbol{r}_1 + (1-\theta)\boldsymbol{r}_2 = (\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2, \theta z_1 + (1-\theta)z_2)$. Check that it is inside the set, that is $|\theta x_1 + (1-\theta)x_2| + |\theta y_1 + (1-\theta)y_2| + |\theta z_1 + (1-\theta)z_2| < 3$. To do this use $|a+b| \le |a| + |b|$, so

$$\begin{aligned} |\theta x_1 + (1-\theta)x_2| + |\theta y_1 + (1-\theta)y_2| + |\theta z_1 + (1-\theta)z_2| &\le \\ \theta(|x_1| + |y_1| + |z_1|) + (1-\theta)(|x_2| + |y_2| + |z_2|) &< 3(\theta + (1-\theta)) = 3. \end{aligned}$$

6. Take two points with radius vectors $\boldsymbol{r}_1 = (x_1, y_1)$ and $\boldsymbol{r}_2 = (x_2, y_2)$. A point between them has a radius vector $\boldsymbol{r} = \theta\boldsymbol{r}_1 + (1-\theta)\boldsymbol{r}_2 = (\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2)$. It is inside the circle if $\|\boldsymbol{r}\| = \sqrt{\boldsymbol{r}\cdot\boldsymbol{r}} \le 1$. Compute it (this is essentially the theorem of cosines):

$$\boldsymbol{r}\cdot\boldsymbol{r} = \theta^2\boldsymbol{r}_1\cdot\boldsymbol{r}_1 + (1-\theta)^2\boldsymbol{r}_2\cdot\boldsymbol{r}_2 + 2\theta(1-\theta)\boldsymbol{r}_1\cdot\boldsymbol{r}_2 \le \theta^2\|\boldsymbol{r}_1\|^2 + (1-\theta)^2\|\boldsymbol{r}_2\|^2 + 2\theta(1-\theta)\|\boldsymbol{r}_1\|\|\boldsymbol{r}_2\|,$$

   as $|\boldsymbol{r}_1\cdot\boldsymbol{r}_2| \le \|\boldsymbol{r}_1\|\|\boldsymbol{r}_2\|$. Continue: the right hand side equals $(\theta\|\boldsymbol{r}_1\| + (1-\theta)\|\boldsymbol{r}_2\|)^2$. Thus, as $\|\boldsymbol{r}_1\|, \|\boldsymbol{r}_2\| \le 1$, one has $\|\boldsymbol{r}\| \le \theta + (1-\theta) = 1$. Done.

7. The set $|x| + |y| + |z| < 3$ has no extreme points, because it is open, so at its each point one can centre a small enough ball, an therefore a short enough line segment, all contained in the set. If the inequality were not strict, the extreme points would be the eight corners – whenever two out of (x, y, z) are zero and one is ± 3.

   For the circle $x^2 + y^2 \le 1$, all the boundary points of the circle are extreme points – by definition.

8. (a) False. Counterexample: take the union of two intersecting lines.

   (b) True. For any pair of points in the intersection, a line segment connecting them belongs to each of the sets, forming the intersection. Then it belongs to the intersection of these sets, by definition of the set intersection.

(c) True. If $\boldsymbol{x}^1 \geq 0$ and $\boldsymbol{x}^2 \geq 0$ (component-wise), then $\theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2 \geq 0$, for $0 \leq \theta \leq 1$.

(d) True. A point $\boldsymbol{y} \in \mathbb{R}^m$ belongs to $\mathcal{C}_A$ if and only if has a pre-image in $\mathcal{C}_n$, namely if there exists $\boldsymbol{x} \in \mathcal{C}_n$, such that $\boldsymbol{y} = A\boldsymbol{x}$. If $\boldsymbol{y}^1, \boldsymbol{y}^2 \in \mathcal{C}_A$, then there exist $\boldsymbol{x}^1, \boldsymbol{x}^2 \in \mathcal{C}_n$, such that $\boldsymbol{y}^k = A\boldsymbol{x}^k$, $k = 1, 2$. Then a point $\theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2 \in \mathcal{C}_n$ (the set $\mathcal{C}_n$ is convex) is a pre-image of $\boldsymbol{y} = \theta \boldsymbol{y}^1 + (1-\theta)\boldsymbol{y}^2$, therefore $\boldsymbol{y} \in \mathcal{C}_A$.

(e) True: $A[\theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2] = \theta A\boldsymbol{x}^1 + (1-\theta)A\boldsymbol{x}^2 = \boldsymbol{b}$.

(f) Convex: Take any $\boldsymbol{y}^1, \boldsymbol{y}^2 \in Y$. Then $\boldsymbol{y}^1 = A\boldsymbol{x}^1$, $y^2 = A\boldsymbol{x}^2$, for some $\boldsymbol{x}^{1,2} \in X$. Any point $\boldsymbol{y}_\theta = \theta \boldsymbol{y}^1 + (1-\theta)\boldsymbol{y}^2$ between $\boldsymbol{y}^1$ and $\boldsymbol{y}^2$ then arises as

$$\boldsymbol{y}_\theta = \theta A\boldsymbol{x}^1 + (1-\theta)A\boldsymbol{x}^2 = A[\theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2] \equiv A\boldsymbol{x}_\theta.$$

As X is convex, $\boldsymbol{x}_\theta \in X$, and therefore $\boldsymbol{y}_\theta \in Y$, because we've found a pre-image $\boldsymbol{x}_\theta \in X$ for it.

Closed – a bit tricky. Take any convergent sequence $\{\boldsymbol{y}^n\}$ in Y, such that $\lim \boldsymbol{y}^n = \boldsymbol{y}$. Y will be closed if $\boldsymbol{y} \in Y$, i.e. for some $\boldsymbol{x}$, $\boldsymbol{y} = A\boldsymbol{x}$. As for each $\boldsymbol{y}^n$ – we know it: $\boldsymbol{y}^n = A\boldsymbol{x}^n$ for some $\boldsymbol{x}^n \in X$. So, if the sequence $\boldsymbol{x}^n$ has a limit point $\boldsymbol{x}$, in X, we will have found a pre-image $\boldsymbol{x}$ for $\boldsymbol{y}$, so $\boldsymbol{y} = A\boldsymbol{x}$ and be done. But what if $\{\boldsymbol{x}^n\}$ has no limit point?

Consider the set K of all $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} = 0$. This set is a linear subspace of $\mathbb{R}^n$, and for every $\boldsymbol{x}^n$ we can write $\boldsymbol{x}^n = \boldsymbol{x}^n_K + \boldsymbol{x}^n_\perp$, where $\boldsymbol{x}^n_K \in K$ and the vector $\boldsymbol{x}^n_\perp$ is normal to K. Then, by linearity, $A\boldsymbol{x}^n = A\boldsymbol{x}^n_\perp$. Let $K^\perp$ be a subspace of $\mathbb{R}^n$, consisting of all vectors orthogonal to K, so each $\boldsymbol{x}^n_\perp \in K^\perp$. Let us denote $S^\perp$ all vectors from $K^\perp$ whose length equals one. Then, as the set $S^\perp$ is closed and bounded, for any $\boldsymbol{d} \in K^\perp$, the length $\|A\boldsymbol{d}\| \geq \epsilon > 0$, for some positive number ϵ, because otherwise would imply that $S^\perp$ contains a vector from K, which cannot be.

Now we can conclude the sequence $\boldsymbol{x}^n_\perp$ is bounded, for otherwise we could find an $\boldsymbol{x}^n_\perp$ of arbitrary large length L, and then the length of $\boldsymbol{y}^n = A\boldsymbol{x}^n_\perp$ would be at least $L\epsilon$, with L going to ∞. But the limit $\boldsymbol{y} = \lim \boldsymbol{x}^n_\perp$ is finite, so it has finite length. Therefore, the sequence $\boldsymbol{x}^n_\perp$, being bounded, must have a limit point $\boldsymbol{x}_\perp$. It is in $X \cap K^\perp$, as X is closed, as well as $K^\perp$, and now $\boldsymbol{y} = A\boldsymbol{x}_\perp$, so Y is closed.