Problem 1. Consider the following linear program:

- Maximize \( 3x_1 + 2x_2 \) subject to:
  - \( x_1 - x_2 \leq -1 \)
  - \( x_1 + x_2 \leq 1 \)
  - \( -2x_1 + x_2 \leq -2 \)
  - \( x_1, x_2 \geq 0 \)

(a) Use the two-phase simplex method on this linear program to prove it’s infeasible, writing out your dictionary and your choices of entering and leaving variables (and why) at each step. (I’d strongly suggest doing this one by hand - practice with the simplex method is a good thing.)

Solution. We start by inserting slack variables and setting up the initial infeasible dictionary, and then adding the extra variable \( x_0 \) for the auxiliary problem.

\[
\begin{align*}
x_3 &= -1 -x_1 + x_2 + x_0 \\
x_4 &= 1 -x_1 - x_2 + x_0 \\
x_5 &= -2 +2x_1 - x_2 + x_0 \\
z' &= -x_0 
\end{align*}
\]

We make our pivot to feasibility, letting \( x_0 \) enter and \( x_5 \) (the most infeasible slack variable) leave.

\[
\begin{align*}
x_0 &= 2 -2x_1 + x_2 + x_5 \\
x_3 &= 1 -3x_1 +2x_2 + x_5 \\
x_4 &= 3 -3x_1 + x_5 \\
z' &= -2 +2x_1 - x_2 - x_5 
\end{align*}
\]

Now we pivot following the standard rule, letting \( x_1 \) enter and \( x_3 \) leave.

\[
\begin{align*}
x_0 &= \frac{4}{3} -\frac{1}{6}x_2 + \frac{2}{3}x_3 + \frac{1}{3}x_5 \\
x_1 &= \frac{4}{3} +\frac{5}{3}x_2 + \frac{2}{3}x_3 + \frac{1}{3}x_5 \\
x_4 &= 2 -2x_2 + x_3 \\
z' &= -\frac{4}{3} +\frac{1}{3}x_2 - \frac{2}{3}x_3 - \frac{1}{3}x_5 
\end{align*}
\]

Following the standard rule again, we let \( x_2 \) enter and \( x_4 \) leave.

\[
\begin{align*}
x_0 &= 1 -\frac{1}{2}x_3 + \frac{1}{6}x_4 + \frac{1}{3}x_5 \\
x_1 &= 1 -\frac{5}{6}x_4 + \frac{3}{3}x_5 \\
x_2 &= 1 +\frac{1}{2}x_3 - \frac{7}{6}x_4 \\
z' &= -1 -\frac{3}{2}x_3 - \frac{6}{6}x_4 - \frac{3}{3}x_5 
\end{align*}
\]

Thus we’ve reached an optimal solution for the auxiliary problem. Since \( z' = -x_0 \) maxes out at \(-1\) (strictly less than 0) we conclude that the original problem was infeasible.

(b) Come up with a linear combination of the first three constraints that give an obviously-false inequality, not involving any variables, which proves again that the linear program is infeasible. (For an example of what I mean, if we had constraints \( x_1 \leq 1 \) and \(-x_1 \leq -2\) then we could take 1 times the first inequality plus 1 times the second to get \( 0 = x_1 + (-x_1) \leq 1 + (-2) \leq -1 \), which is false). Your coefficients may look familiar from part (a) - we’ll see an explanation for these “magic coefficients” a bit later on in the course.
Solution. One way to do this is to take 1/2 times the first inequality plus 1/6 times the second plus 1/3 times the third:

\[ 0 = \frac{1}{2}(x_1 - x_2) + \frac{1}{6}(x_1 + x_2) + \frac{1}{3}(-2x_1 + x_2) \leq \frac{1}{2}(-1) + \frac{1}{6}(1) + \frac{1}{3}(-2) = -1. \]

The coefficients here are exactly the negatives of the ones that show up in the expression for \( z' \) in our final dictionary! This will be explained later on as having to do with duality theory.

Problem 2. Consider the Chvátal’s example of a dictionary that cycles under the standard rule that we looked at in class:

\[
\begin{align*}
x_5 &= -0.5x_1 + 5.5x_2 + 2.5x_3 - 9x_4 \\
x_6 &= -0.5x_1 + 1.5x_2 + 0.5x_3 - x_4 \\
x_7 &= 1 - x_1 \\
z &= 10x_1 - 57x_2 - 9x_3 - 24x_4.
\end{align*}
\]

Perform the simplex method but following Bland’s rule rather than the standard rule, and solve the given LP. At what step do the rules first start to differ? (Feel free to use a computer to help with the computations, but for each step write down which variables you choose to enter and exit and why, and the resulting dictionary).

Solution. Bland’s rule is that we always choose the entering variable to have the smallest subscript among all of those with positive coefficients in \( z \), and the leaving variable having the smallest subscript among the ones that preserve feasibility (note this means the leaving variable is chosen the same way as the standard rule). For our initial dictionary, we choose \( x_1 \) to enter and \( x_5 \) to leave, and get

\[
\begin{align*}
x_1 &= 11x_2 + 5x_3 - 18x_4 - 2x_5 \\
x_6 &= -4x_2 - 2x_3 + 8x_4 + x_5 \\
x_7 &= 1 - 11x_2 - 5x_3 + 18x_4 + 2x_5 \\
z &= 53x_2 + 41x_3 - 204x_4 - 20x_5.
\end{align*}
\]

We next choose \( x_2 \) to enter and \( x_6 \) to leave.

\[
\begin{align*}
x_1 &= -0.5x_3 + 2x_4 + 0.25x_5 - 0.25x_6 \\
x_2 &= -0.5x_3 + 4x_4 + 0.75x_5 - 2.75x_6 \\
x_7 &= 1 + 0.5x_3 - 4x_4 - 0.75x_5 - 13.25x_6 \\
z &= 14.5x_3 - 98x_4 - 6.75x_5 - 13.25x_6.
\end{align*}
\]

For the next pivot we have \( x_3 \) enter and \( x_1 \) leave.

\[
\begin{align*}
x_2 &= x_1 - 2x_4 - 0.5x_5 + 2.5x_6 \\
x_3 &= -2x_1 + 8x_4 + 1.5x_5 - 5.5x_6 \\
x_7 &= 1 - x_1 \\
z &= -29x_1 + 18x_4 + 15x_5 - 93x_6.
\end{align*}
\]

We choose \( x_4 \) to enter and \( x_2 \) to leave.

\[
\begin{align*}
x_3 &= 2x_1 - 4x_2 - 0.5x_5 + 4.5x_6 \\
x_4 &= 0.5x_1 - 0.5x_2 - 0.25x_5 - 1.25x_6 \\
x_7 &= 1 - x_1 \\
z &= -20x_1 - 9x_2 + 10.5x_5 - 70.5x_6.
\end{align*}
\]

We choose \( x_5 \) to enter and \( x_3 \) to leave.

\[
\begin{align*}
x_4 &= -0.5x_1 + 1.5x_2 + 0.5x_3 - x_6 \\
x_5 &= 4x_1 - 8x_2 + 2x_3 + 9x_6 \\
x_7 &= 1 - x_1 \\
z &= 22x_1 - 93x_2 - 21x_3 + 24x_6.
\end{align*}
\]
We then choose $x_1$ to enter and $x_4$ to leave (which is the place where Bland’s rule starts to differ from the standard rule!)

\[
\begin{align*}
    x_1 &= 3x_2 + x_3 - 2x_4 - 2x_6 \\
    x_5 &= 4x_2 + 2x_3 - 8x_4 + 1x_6 \\
    x_7 &= 1 - 3x_2 - x_3 + 2x_4 + 2x_6 \\
    z &= -27x_2 + x_3 - 44x_4 - 20x_6
\end{align*}
\]

We then choose $x_3$ to enter and $x_7$ to leave, which (finally) is a nondegenerate pivot and brings us to the optimal dictionary.

\[
\begin{align*}
    x_1 &= 1 \\
    x_4 &= 1 - 3x_2 + 2x_4 + 2x_6 - x_7 \\
    x_5 &= 2 - 2x_2 - 4x_4 + 5x_6 - 2x_7 \\
    z &= 1 - 30x_2 - 42x_4 - 18x_6 - x_7
\end{align*}
\]

**Problem 3.** Take the final (optimal) dictionary you got from Problem 2, and write out all of the pieces that show up in the formula for the dictionary in matrix notation (as discussed in class on Monday): the vectors of variables $\vec{x}_B$ and $\vec{x}_N$, the matrices $A_B$ and $A_N$, and the vectors $\vec{b}$, $\vec{c}_B$ and $\vec{c}_N$. Compute $A_B^{-1}\vec{b}$ and confirm this is the vector of constants in your dictionary, and compute $\vec{c}^\top A_B^{-1}\vec{b}$ and confirm this is the constant in your final expression of $z$. (This partially checks that our formulas do actually recover the final dictionary).

**Solution.** From the last dictionary, $x_1, x_3, x_5$ are the basic variables, so we take

\[
\vec{x}_B = \begin{bmatrix} x_1 \\ x_3 \\ x_5 \end{bmatrix} \quad \vec{x}_N = \begin{bmatrix} x_2 \\ x_4 \\ x_6 \\ x_7 \end{bmatrix}.
\]

We get $A_B$ and $A_N$ by splitting up the augmented matrix $[A|I]$ as appropriate, where $A$ is the original matrix of coefficients (which is the *negative* of the coefficients from the slack variables!)

\[
[A|I] = \begin{bmatrix}
0.5 & -5.5 & -2.5 & 9 & 1 & 0 & 0 \\
0.5 & -1.5 & -0.5 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
A_B = \begin{bmatrix}
0.5 & -2.5 & 1 \\
0.5 & -0.5 & 0 \\
1 & 0 & 0
\end{bmatrix} \quad A_N = \begin{bmatrix}
-5.5 & 9 & 0 & 0 \\
-1.5 & 1 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}.
\]

Similarly we can read off $\vec{b}$ and the original vector $\vec{c}$ from the first dictionary, which we augment and split up into $\vec{c}_B$ and $\vec{c}_N$ according to the choice of basic variables:

\[
\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \vec{c} = \begin{bmatrix} 10 \\ -57 \\ -9 \\ -24 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \vec{c}_B = \begin{bmatrix} 10 \\ -9 \\ 0 \end{bmatrix} \quad \vec{c}_N = \begin{bmatrix} -57 \\ -24 \\ 0 \\ 0 \end{bmatrix}.
\]

Finally, we compute $A_B^{-1}$ using any method we want to get the inverse of a matrix (a calculator or computer is fine):

\[
A_B^{-1} = \begin{bmatrix}
0 & 0 & 1 \\
0 & -2 & 1 \\
1 & -5 & 2
\end{bmatrix}.
\]
Then since $\vec{b}$ is the third elementary basis vector, $A_{B}^{-1}\vec{b}$ is just the third column of this matrix:

$$A_{B}^{-1}\vec{b} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

which is indeed the vector of constant terms in the constraints for our final dictionary. Finally,

$$\vec{c}^T A_{B}^{-1}\vec{b} = \begin{bmatrix} 1 & 10 & -9 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 10 - 9 + 0 = 1$$

which is the constant term in $z$ in our final dictionary.

**Problem 4.** Suppose that you start with a nondegenerate dictionary (all of the constant terms for the constraints are positive) and that after you pick the entering variable via the simplex method, there is not a tie for which variable should leave (so we don’t need to use the “smallest subscript” tiebreaker part of the standard rule). Is it possible for the resulting dictionary to be degenerate? Explain why or why not.

**Solution.** It cannot be degenerate; we can justify this by thinking about where the constants in the new dictionary come from. Suppose $x_i$ entered the dictionary and $x_j$ left. For the variable $x_i$ that just entered the dictionary, the constant is $b_j/(-a_{ij})$ where $b_j$ was the constant for the equation for $x_j$ in the original dictionary, and $a_{ij}$ was the (negative) coefficient of $x_i$ in that dictionary. By assumption the original dictionary is nondegenerate, $b_j > 0$, so the new constant $b_j/(-a_{ij})$ is also strictly positive.

Then look at the coefficients for the other variables $x_k$ that remained in the dictionary from the previous step. If they started with a constant $b_k$ and coefficient $a_{ik}$ for $x_i$, then after substituting in our new equation for $x_i$ the new constant will be

$$b_k + a_{ik} \frac{b_j}{-a_{ij}}.$$ 

Our assumption that the original dictionary was nondegenerate tells us that $b_j > 0$. Then, go back to how the exiting variable was chosen to be $x_j$: it was picked because we raised the value of $x_i$ as much as possible, and found the first basic variable to hit zero was $x_j$, which would have happened when $x_i = b_j/(-a_{ij})$. Moreover our assumption that there was not a tie means that no other variable hit zero at the same time - so by definition substituting in $x_i = b_j/(-a_{ij})$ to our formula for $x_k$ must still give a positive value. But making this substitution gives our new constant $b_k + a_{ik} \frac{b_j}{-a_{ij}}$ is strictly positive.

**Problem 5.** Suppose we have a linear program in standard form, which we write in matrix notation:

maximize $\vec{c} \cdot \vec{x}$ subject to $A\vec{x} \leq \vec{b}$ and $\vec{x} \geq 0$ (where we fix a $m \times n$ matrix $A$ and two vectors $\vec{b} \in \mathbb{R}^n$ and $\vec{c} \in \mathbb{R}^m$, and treat $\vec{x}$ as a variable in $\mathbb{R}^n$).

**a)** Prove that if $\vec{x}_1$ and $\vec{x}_2$ are two feasible solutions for our constraint, then so is any weighted average of them:

$$t\vec{x}_1 + (1-t)\vec{x}_2 \quad 0 \leq t \leq 1.$$ 

The set of all of these weighted averages (for all values of $t$ in $[0, 1]$) forms a line segment between $\vec{x}_1$ and $\vec{x}_2$. What you’ve proven says that anytime we have two points in our feasible region, the line segment between them is contained in the feasible region too; this is the definition of the feasible region being convex. (Note: You can prove this using algebraic manipulations, but make sure to justify why those manipulations are valid - especially ones involving inequalities of vectors!)
Solution. $x_1$ and $x_2$ being feasible means that we have $x_1, x_2 \geq 0$ and $Ax_1, Ax_2 \leq b$. Note that we can multiply an inequality of vectors by a nonnegative coefficient and keep the same inequality (because we can do this for coordinatewise for the real numbers we’re comparing); since $0 \leq t \leq 1$ we have $t, (1-t) \geq 0$ and thus we can obtain all of the inequalities

$$tx_1 \geq 0 \quad (1-t)x_2 \geq 0 \quad A(tx_1) = t(Ax_1) \leq tb \quad A((1-t)x_1) = (1-t)(Ax_1) \leq (1-t)b$$

(using that scalar multiplication can be brought inside of matrix multiplication). Next, we note we can add together two inequalities of vectors (of the same size) and get another inequality of vectors, again because we can do this coordinatewise for the real numbers. Adding together the first two inequalities above gives

$$tx_1 + (1-t)x_2 \geq 0,$$

the first thing needed for this weighed average to be feasible. Adding together the last two and using that matrix multiplication distributes over addition we have

$$A(tx_1 + (1-t)x_2) = A(tx_1) + A((1-t)x_1) \leq tb + (1-t)b = b,$$

which is the other thing needed to conclude it’s a feasible solution.

(b) Suppose that $x_1$ and $x_2$ are both optimal solutions to the given LP. Prove that any weighted average $tx_1 + (1-t)x_2$ is also an optimal solution.

Solution. If both are optimal, then their values at the objective function $c \cdot x_1$ and $c \cdot x_2$ are equal to each other and to the optimum value $z_{max}$. Then, we can simply compute the value of the objective function at the weighted average, using basic rules for manipulating dot products:

$$c \cdot (tx_1 + (1-t)x_2) = tc \cdot x_1 + tc \cdot x_2 = tz_{max} + (1-t)z_{max} = z_{max},$$

which is again equal to the optimum value. (And we proved that the weighted average is a feasible solution in part (a)!)