The Structure of Polyhedra & Linear Programming

John MacLaren Walsh, Ph.D.

April 2, 2014 & April 9, 2014

1 References

Numerous books have been written about linear programming. I recommend the first one below for a rigorous graduate level treatment that can at times be a tough read, and the second one below for a more intuitive approach.

- M. Padberg, **Linear Optimization and Extensions, 2nd Ed.** Springer-Verlag, 1999.
- V. Chvátal, **Linear Programming**, W. H. Freeman, 1983.

2 The Structure of Vector Subspaces and Affine Subspaces

The motivation for the first part of this lecture is a couple of facts which you came to know very well when you took linear algebra. A vector subspace $\mathcal{V} \subseteq \mathbb{R}^N$ may be represented in (at least) two forms. Let the dimension of $\mathcal{V}$ be d. In the first form, we think of $\mathcal{V}$ as the set of vectors obeying $N-d$ equalities

$$\mathcal{V} = \left\{ \mathbf{x} \in \mathbb{R}^N \,|\mathbf{A}\mathbf{x} = \mathbf{0} \right\} \tag{1}$$

for an appropriately selected matrix $\mathbf{A} \in \mathbb{R}^{(N-d)\times N}$.

In the second form, we think of $\mathcal{V}$ as the set of linear combinations (span) of a collection of d (linearly independent) basis vectors

$$\mathcal{V} = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_d\} = \left\{ \sum_{i=1}^{d} \alpha_i \mathbf{v}_i \mid \alpha_i \in \mathbb{R} \ \forall i \in \{1, \dots, d\} \right\} \tag{2}$$

The next section of this lecture extends this idea of dual representation to polyhedra, which contain vector subspaces as special cases.

3 The Structure of Polyhedra (Padberg, Ch. 7)

3.1 The Linear (Inequality) Description of a Polyhedron

A familiar definition of a *polyhedron*, $\mathcal{P}$, is the set of $\mathbf{x} \in \mathbb{R}^N$ obeying a system of linear inequalities, i.e.

$$\mathcal{P} := \left\{ \mathbf{x} \in \mathbb{R}^N \mid \mathbf{H}\mathbf{x} \leq \mathbf{h} \right\} \tag{3}$$

where $\mathbf{H} \in \mathbb{R}^{M\times N}$, $\mathbf{h} \in \mathbb{R}^M$, and the inequality is understood to hold elementwise. Note that it suffices to consider inequalities of the form $\leq$ here because, were we to have a $\mathbf{a}^T\mathbf{x} \geq b$ we can negate it to get an equivalent $-\mathbf{a}^T\mathbf{x} \leq -b$ and were we to have an $\mathbf{a}^T\mathbf{x} = b$ we can use the pair of inequalities $\mathbf{a}^T\mathbf{x} \leq b$, $-\mathbf{a}^T\mathbf{x} \leq -b$.

We drew three different examples of polyhedra in $\mathbf{R}^2$.

3.2 Lines, Extreme Points, and Extreme Directions

Consider the polyhedron as defined in (3). Observe that if there exists some $\mathbf{y} \neq \mathbf{0}$ for which $\mathbf{H}\mathbf{y} = \mathbf{0}$, and if we take any point $\mathbf{x} \in \mathcal{P}$, then we will also have $(\mathbf{x} + \alpha\mathbf{y}) \in \mathcal{P}$ for any $\alpha \in \mathbb{R}$. That is, the line $\{\mathbf{x} + \alpha\mathbf{y} \,|\alpha \in \mathbb{R}\}$ lies in $\mathcal{P}$. If we had two different such $\mathbf{y}$s, we could add any linear combination of them to any point $\mathbf{x} \in \mathcal{P}$ and get another point in $\mathcal{P}$, and so on. Formalizing this idea, the *lineality space* of the polyhedron is the vector space of solutions $\mathcal{L}_{\mathcal{P}} = \left\{ \mathbf{y} \in \mathbb{R}^N \middle|\ \mathbf{H}\mathbf{y} = \mathbf{0} \right\}$.

Clearly, the lineality space of the polyhedron is non-empty if and only if rank($\mathbf{H}$) $< N$. In this case we can select a basis $\mathbf{g}_1, \ldots, \mathbf{g}_I$ for the null space of $\mathbf{H}$, and represent $\mathcal{L}_{\mathcal{P}} = \text{span}\{\mathbf{g}_1, \ldots, \mathbf{g}_I\}$. Here, we have $I = N - \text{rank}(\mathbf{H})$. Collecting these basis vectors into the matrix $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_I]^T$, then the orthogonal complement of $\mathcal{L}_{\mathcal{P}}$, denoted $\mathcal{L}_{\mathcal{P}}^{\perp}$ can be represented as $\mathcal{L}_{\mathcal{P}}^{\perp} = \{\mathbf{z} | \mathbf{G}\mathbf{z} = \mathbf{0}\}$.

We can remove the lines from $\mathcal{P}$, creating a line free polyhedron, by intersecting it with $\mathcal{L}_{\mathcal{P}}^{\perp}$ to get

$$\mathcal{P}^0 = \mathcal{P} \cap \mathcal{L}_{\mathcal{P}}^{\perp} = \left\{\mathbf{x} \in \mathbb{R}^N \mid \mathbf{H}\mathbf{x} \leq \mathbf{h}, \mathbf{G}\mathbf{x} = \mathbf{0} \ \right\} \tag{4}$$

Furthermore, each point $\mathbf{x} \in \mathcal{P}$ now has a unique representation as $\mathbf{x} = \mathbf{z} + \mathbf{y}$ with $\mathbf{y} \in \mathcal{L}_{\mathcal{P}}$ and $\mathbf{z} \in \mathcal{P}^0$. That is, we have the alternate representation of a polyhedron as

$$\mathcal{P} = \mathcal{P}^0 + \mathcal{L}_{\mathcal{P}} \tag{5}$$

Now that we have removed the lines from $\mathcal{P}$ to get $\mathcal{P}^0$, let's continue on to consider the properties of non-empty line free polyhedra.

It turns out that although we have removed the lines, we still have the possibility for unboundedness. To see this observe that if we have $\mathbf{y}$ such that $\mathbf{H}\mathbf{y} \leq \mathbf{0}$, then for any $\mathbf{x} \in \mathcal{P}$, we will have the halfline $(\mathbf{x} + \mu\mathbf{y}) \in \mathcal{P}$, any $\mu \geq 0$. For this reason we are interested in the *asymptotic cone* / *recession cone* / *characteristic cone* of a polyhedral set, which is defined as

$$\mathcal{C}_{\infty}(\mathbf{H}) = \left\{\mathbf{y} \in \mathbb{R}^N \,|\mathbf{H}\mathbf{y} \leq \mathbf{0}\right\} \tag{6}$$

Observe that if $\mathbf{y}_1$ and $\mathbf{y}_2$ are in $\mathcal{C}_{\infty}$, then we must also have $(\mu_1\mathbf{y}_1 + \mu_2\mathbf{y}_2) \in \mathcal{C}_{\infty}$ for any $\mu_1, \mu_2 \geq 0$! Getting rid of the lines again to get

$$\mathbf{C}_{\infty}^0 = \mathbf{C}_{\infty} \cap \mathcal{L}_{\mathcal{P}}^{\perp} \tag{7}$$

We will define an *extreme ray* of $\mathcal{C}_{\infty}^0$ (resp. a *extreme direction* of $\mathcal{P}^0$) to be a $\{\gamma\mathbf{y}_* | \gamma \geq 0\} \in \mathcal{C}_{\infty}^0$ such that whenever

$$\{\gamma\mathbf{y}_* | \gamma \geq 0\} = \mu_1\{\gamma\mathbf{y}_1 | \gamma \geq 0\} + \mu_2\{\gamma\mathbf{y}_2 | \gamma \geq 0\} \tag{8}$$

for some $\mu_1, \mu_2 > 0$ and $\{\gamma\mathbf{y}_1 | \gamma \geq 0\} \in \mathcal{C}_{\infty}^0$ and $\{\gamma\mathbf{y}_2 | \gamma \geq 0\} \in \mathcal{C}_{\infty}^0$ we must have $\{\gamma\mathbf{y}_1 | \gamma \geq 0\} = \{\gamma\mathbf{y}_2 | \gamma \geq 0\} = \{\gamma\mathbf{y}_* | \gamma \geq 0\}$. This will happen if and only if $\mathbf{y}_*$ obeys exactly $N-1$ linearly independent rows of $\mathbf{H}\mathbf{y} \leq \mathbf{0}$ with equality (and any remaining inequalities linearly independent of these inequalities with strict inequality). From this property it is clear that a polyhedron can have at most a finite number of extreme directions.

An *extreme point* of $\mathcal{P}^0$ is a point which may not be represented as a convex combination of any two points other than itself. That is, if $\mathbf{x}$ is an extreme point of $\mathcal{P}$ and if $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{P}$ with $\lambda \in (0,1)$ then we must have $\mathbf{x} = \mathbf{x}_1 = \mathbf{x}_2$. $\mathbf{x}_0$ is an extreme point of the polyhedron $\mathcal{P}$ if and only if $\mathbf{H}\mathbf{x}_0 \leq \mathbf{h}$ and if there exists a $N \times N$ submatrix $\mathbf{H}_{\mathcal{B}}$ of $\mathbf{H}$, that is full rank (N) such that $\mathbf{H}_{\mathcal{B}}\mathbf{x} = \mathbf{h}_{\mathcal{B}}$. From this fact, it is clear that a polyhedron can have only at most a finite number of extreme points.

Bounded polyhedra are called *polytopes*. Polytopes have no extreme rays and the trivial lineality space.

3.3 A Second Description of a Polyhedron

Polyhedra have a second, alternate definition/representation. Namely, a polyhedron has a second representation

$$\mathcal{P} = \mathcal{L}_{\mathcal{P}} + \text{conv}\mathcal{S} + \text{cone}\mathcal{T} \tag{9}$$

Here $\mathcal{L}_{\mathcal{P}}$ is the *lineality space* of the polyhedron, $\mathcal{S}$ is the finite set of *extreme points* of the polyhedron $\mathcal{P}^0$, and $\mathcal{T}$ is the finite set of *extreme directions* (scaled to unit length) of the polyhedron $\mathcal{P}^0$ (the extreme rays of the recession cone $\mathcal{C}_{\infty}^0$).

Put another way, under this second definition, any point $\mathbf{x} \in \mathcal{P}$ can be represented as

$$\mathbf{x} = \sum_{i=1}^{I} \alpha_i \mathbf{g}_i + \sum_{j=1}^{J} \beta_j \mathbf{s}_j + \sum_{k=1}^{K} \gamma_k \mathbf{t}_k \tag{10}$$

where I is the dimension of the lineality space $\mathcal{L}_{\mathcal{P}}$, and $\text{span}\{\mathbf{g}_1, \ldots, \mathbf{g}_I\} = \mathcal{L}_{\mathcal{P}}$, J is the number of extreme points, K is the number of extreme directions, $\alpha_i \in \mathbb{R} \forall i$, $\beta_j \in \mathbb{R}, \beta_j \geq 0, \ \forall j, \sum_{j=1}^{J} \beta_j = 1$, and $\gamma_k \in \mathbb{R}, \ \gamma_k \geq 0$.

4 Linear Programs

4.1 Properties of their Solutions

Next, we discussed the linear program of the form

$$\min_{\mathbf{x}\in\mathcal{P}} \mathbf{c}\cdot\mathbf{x} \tag{11}$$

with $\mathcal{P}$ defined as in (3). The second representation of polyhedra introduced in (10) has significant implications for the structure of solutions to linear programs. To see this, we observed that since every point $\mathbf{x}\in\mathcal{P}$ could be written as (3), we could rewrite the problem (11) alternatively as

$$\min_{\boldsymbol{\alpha}\in\mathbb{R}^I,\ \boldsymbol{\beta}\geq\mathbf{0},\ \mathbf{1}^T\boldsymbol{\beta}=1,\ \boldsymbol{\gamma}\geq\mathbf{0}} \mathbf{c}\cdot\left(\sum_{i=1}^{I}\alpha_i\mathbf{g}_i+\sum_{j=1}^{J}\beta_j\mathbf{s}_j+\sum_{k=1}^{K}\gamma_k\mathbf{t}_k\right) = \min_{\boldsymbol{\alpha}\in\mathbb{R}^I,\ \boldsymbol{\beta}\geq\mathbf{0},\ \mathbf{1}^T\boldsymbol{\beta}=1,\ \boldsymbol{\gamma}\geq\mathbf{0}} \sum_{i=1}^{I}\alpha_i\mathbf{c}\cdot\mathbf{g}_i+\sum_{j=1}^{J}\beta_j\mathbf{c}\cdot\mathbf{s}_j+\sum_{k=1}^{K}\gamma_k\mathbf{c}\cdot\mathbf{t}_k \tag{12}$$

From this we observe

- If for any $i\in\{1,\ldots,I\}$, $\mathbf{c}\cdot\mathbf{g}_i\neq 0$, the linear program is unbounded below!
- If for any $k\in\{1,\ldots,K\}$, $\mathbf{c}\cdot\mathbf{t}_k<0$, the linear program is unbounded below!
- If for all $i\in\{1,\ldots,I\}$, $\mathbf{c}\cdot\mathbf{g}_i=0$ and for all $k\in\{1,\ldots,K\}$ $\mathbf{c}\cdot\mathbf{t}_k\geq 0$, then the solution to the linear program is (set $\gamma_k=0$)

$$\min_{\boldsymbol{\beta}\geq\mathbf{0},\ \mathbf{1}^T\boldsymbol{\beta}=1,} \sum_{j=1}^{J}\beta_j\mathbf{c}\cdot\mathbf{s}_j = \min_{j\in\{1,\ldots,J\}} \mathbf{c}\cdot\mathbf{s}_j \tag{13}$$

  and is attained for instance at $\mathbf{x}=\mathbf{s}_j$, $j\in\arg\min_{j\in\{1,\ldots,J\}}\mathbf{c}\cdot\mathbf{s}_j$.

From this we observe that the solution to *any* bounded linear program over the polyhedron must be attained at an extreme point. The algorithm we will discuss next, the simplex method, will exploit this fact in order to find the solution to the linear program in an efficient manner.

4.2 Standard Form (Chapter 2)

The most general form of a linear program could be considered to be

$$\min_{\left\{\mathbf{x}\,\middle|\,\begin{array}{c}\mathbf{A}\mathbf{x}\leq\mathbf{b}\\ \mathbf{A}'\mathbf{x}=\mathbf{b}'\end{array}\right\}} \mathbf{c}\cdot\mathbf{x}$$

All linear programs can be brought into *standard form*, in which all of the constraints are equalities except for positivity constraints on the variables

$$\min_{\left\{\mathbf{x}\,\middle|\,\begin{array}{c}\mathbf{A}\mathbf{x}=\mathbf{b}\\ \mathbf{x}\geq\mathbf{0}\end{array}\right\}} \mathbf{c}\cdot\mathbf{x}$$

This is performed by first replacing any variable x_i which is unrestricted in sign with two which must be positive $x_i=x_i^+-x_i^-$ and by introducing new slack variables indicating the slack in each inequality. For instance, if after introducing our sign restricted variables into the general form we have N variables, then the ith inequality $\mathbf{a}_i\cdot\mathbf{x}\leq b_i$ can be replaced by the equality $\mathbf{a}_i\cdot\mathbf{x}+x_{N+i}=b_i$ where the slack variable $x_{N+i}\geq 0$.

4/7/2014 lecture ended here.

Additionally, surprisingly it suffices to assume that the matrix $\mathbf{A}$ in standard form is full rank, since were it not to be full rank, we can transform the problem again according to the simultaneous constraints

$$\begin{array}{c}\mathbf{A}\mathbf{x}\leq\mathbf{b}\\ \mathbf{A}\mathbf{x}\geq\mathbf{b}\end{array} \leftrightarrow \begin{pmatrix}\mathbf{A} & \mathbf{I} & \mathbf{0}\\ -\mathbf{A} & \mathbf{0} & \mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{x}\\ \mathbf{s}\\ \mathbf{t}\end{pmatrix}=\begin{pmatrix}\mathbf{b}\\ -\mathbf{b}\end{pmatrix}$$

with $\mathbf{x} \geq \mathbf{0},\ \mathbf{s} \geq \mathbf{0}, \mathbf{t} \geq \mathbf{0}$. This new matrix (via its structure) has full row rank regardless of the rank of the original $\mathbf{A}$, and the associated problem is still in standard form. Thus it suffices to assume that a linear program is in standard form with $\mathbf{A}$ full row rank.

Observe that the transformation into standard form removes any lineality space by the possible incorporation of new variables, because after the transformation into standard form, the matrix defining the polyhedron is full rank (this is already evident from the constraint $\mathbf{x} \geq \mathbf{0}$ alone).

4.3 Some Other Linear Programming Jargon & Facts

Start with a linear program in standard form with $\mathbf{A}$ a $M \times N$ matrix ($M \leq N$) of rank M.

- A *basis* is a collection of M linearly independent columns from $\mathbf{A}$ with indices in the set $\mathcal{C}$ collected into a matrix $\mathbf{B} := \mathbf{A}_{\mathcal{C}}$.
- A basis $\mathbf{B}$ is a *feasible basis* if $\mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}$.
- A *feasible solution* $\mathbf{x}$ is one that obeys $\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0$.
- A *basic feasible solution* is one for which the columns of $\mathbf{A}$ corresponding to the non-zero entries of $\mathbf{x}$ (denoted by $\mathbf{A}_{\mathcal{I}_x}$ and $\mathcal{I}_x \subset \{1, \dots N\}$ respectively) are linearly independent. Note that this means that the number of non-zero entries in a basic feasible solution must be $\leq M$.
- We say a basic feasible solution is a *degenerate basic feasible solution* if the $\text{rank}(\mathbf{A}_{\mathcal{I}_x}) < M$.

Note that any feasible basis gives a basic feasible solution $\mathbf{x}_{\mathcal{C}} = \mathbf{B}^{-1}\mathbf{b},\ \mathbf{x}_{\mathcal{N}\setminus\mathcal{C}} = \mathbf{0}$ via

$$\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}_{\mathcal{C}} + \mathbf{A}_{\mathcal{N}\setminus\mathcal{C}}\mathbf{x}_{\mathcal{N}\setminus\mathcal{C}} = \mathbf{b} \implies \mathbf{x}_{\mathcal{C}} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{C}}\mathbf{x}_{\mathcal{N}\setminus\mathcal{C}}. \tag{14}$$

Similarly, any basic feasible solution can be associated with at least one feasible basis containing $\mathbf{A}_{\mathcal{I}_x}$ and, if necessary, some extra columns from $\mathbf{A}_{\mathcal{N}\setminus\mathcal{I}_x}$. (The extra columns are necessary if and only if the feasible basis was degenerate.)

Note the conditions for $\mathbf{x}$ to be a *basic feasible solution* are equivalent to the conditions for it to be an *extreme point* of the polyhedral constraint (feasible) set $\mathcal{X} := \{\mathbf{x} \,|\, \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$. To see this, recall that for it to be an extreme point, there needed a subset of N linearly independent inequalities which $\mathbf{x}$ satisfied with equality in addition to lying within the set (i.e. obeying the other inequalities). For a basic feasible solution, equality is obtained in the inequalities $\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x}_{\mathcal{N}\setminus\mathcal{I}_\mathbf{x}} \geq \mathbf{0}$, and there are more than N linearly independent inequalities among these by the requirement that the columns of $\mathbf{A}$ given by $\mathbf{A}_{\mathcal{I}_x}$ be linearly independent (see (15) below).

Next, we must note that although every extreme point is associated with at least one feasible basis, it is possible that the same extreme point is defined by multiple feasible bases. This occurs when the associated basic feasible solution is degenerate.

Every extreme point (i.e. basic feasible solution) of $\mathcal{X}$ has an objective function $\mathbf{c} \cdot \mathbf{x}$ which it is the unique optimal solution for. This follows by selecting $\mathbf{c}_{\mathcal{I}_x} = 0$ and $\mathbf{c}_{\mathcal{N}\setminus\mathcal{I}_x} = 1$, because the system

$$\begin{pmatrix} \mathbf{A}_{\mathcal{I}_x} & \mathbf{A}_{\mathcal{N}\setminus\mathcal{I}_x} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{\mathcal{I}_x} \\ \mathbf{x}_{\mathcal{N}\setminus\mathcal{I}_x} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} \tag{15}$$

has a unique solution, and $\mathbf{c} \cdot \mathbf{x}'$ is the sum of the (non-negative) elements of $\mathbf{x}'_{\mathcal{N}\setminus\mathcal{I}_x}$, which is thus ≥ 0, in which equality is attained for a feasible $\mathbf{x}'$ iff $\mathbf{x}' = \mathbf{x}$.

4.4 Detecting Optimality

Let $\mathbf{B}$ be a feasible basis, and let $\mathcal{B}$ be the corresponding column indices from $\mathbf{A}$. If $\mathbf{c}^T_{\mathcal{N}\setminus\mathcal{B}} - \mathbf{c}^T_{\mathcal{B}}\mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{B}} \geq \mathbf{0}$ then the basic feasible solution $\mathbf{x}_{\mathcal{B}} = \mathbf{B}^{-1}\mathbf{b}$ with $\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}} = \mathbf{0}$ is optimal! Indeed, for such a basis $\mathbf{B}$, and for any feasible $\mathbf{x}$, we have $\mathbf{x}_{\mathcal{B}} + \mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{B}}\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}} = \mathbf{B}^{-1}\mathbf{b}$, and thus for any feasible $\mathbf{x}$ we have

$$\begin{aligned} \mathbf{c} \cdot \mathbf{x} &= \mathbf{c}^T_{\mathcal{B}}\left(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{B}}\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}}\right) + \mathbf{c}^T_{\mathcal{N}\setminus\mathcal{B}}\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}} \quad (16) \\ &= \mathbf{c}^T_{\mathcal{B}}\mathbf{B}^{-1}\mathbf{b} + \left(\mathbf{c}^T_{\mathcal{N}\setminus\mathcal{B}} - \mathbf{c}^T_{\mathcal{B}}\mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{B}}\right)\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}} \quad (17) \\ &\geq \mathbf{c}^T_{\mathcal{B}}\mathbf{B}^{-1}\mathbf{b} \quad (18) \end{aligned}$$

since $\mathbf{x}_{\mathcal{N}\setminus\mathcal{B}} \geq \mathbf{0}$.

Note this is a sufficient, but not necessary, condition for optimality. However, if $\mathbf{x}$ is a basic feasible nondegenerate solution that is optimal, then $\mathbf{c}_{\mathcal{N}\setminus\mathcal{B}} - \mathbf{c}_{\mathcal{B}}\mathbf{B}^{-1}\mathbf{A}_{\mathcal{N}\setminus\mathcal{B}} \geq \mathbf{0}$.

4.5 Detecting Unboundedness

If there is a feasible basis $\mathbf{B} = \mathbf{A}_{\mathcal{B}}$ such that there is a non-basic (i.e. $j \in \mathcal{N} \setminus \mathcal{B}$) index such that $c_j - \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1} \mathbf{a}_j < 0$ and $\mathbf{B}^{-1} \mathbf{a}_j \leq 0$ then the objective function is not bounded below. This can be seen by choosing $\mathbf{x}_{\mathcal{B}}(\lambda) = \mathbf{B}^{-1}\mathbf{b} - \lambda \mathbf{B}^{-1} \mathbf{a}_j$, $x_j := \lambda$, and $\mathbf{x}_{\mathcal{N}-\mathcal{B}-\{j\}} = \mathbf{0}$ since this is $\geq \mathbf{0}$ for all $\lambda \geq 0$, and is such that $\mathbf{A}\mathbf{x}(\lambda) = \mathbf{b}$, but has objective value $\mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1} \mathbf{b} + \lambda (c_j - \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1} \mathbf{a}_j)$ which $\to -\infty$ as $\lambda \to \infty$.

4.6 Changing the Basis: Pivoting

Suppose we are working with a feasible basis $\mathbf{B} := \mathbf{A}_{\mathcal{B}}$ and the associated basic feasible solution, for which neither the criteria for detecting optimality and unboundedness are satisfied. In other words, there is a $j \in \mathcal{N} \setminus \mathcal{B}$ such that

$$c_j - \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1} \mathbf{a}_j < 0 \tag{19}$$

Then, if we replace the a well selected column in $\mathbf{B}$ with the vector $\mathbf{a}_j$ we obtain another feasible basis $\mathbf{B}' := \mathbf{A}_{\mathcal{B}'}$ and corresponding basic feasible solution $\mathbf{x}'$ with a lower or equal cost $\mathbf{c} \cdot \mathbf{x}' \leq \mathbf{c} \cdot \mathbf{x}$. The column to remove is selected according to the minimal ratio

$$r \in \arg \min_{r | [\mathbf{B}^{-1}\mathbf{a}_j]_r > 0} \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} \tag{20}$$

Of course, this means that we must have $\mathbf{B}^{-1}\mathbf{a}_j \not\leq \mathbf{0}$ (so that it has at least one positive element), but this is fine because were this not to be the case we would know the linear program is unbounded below.

Letting the rth index in $\mathcal{B}$ (i.e. the column index of $\mathbf{A}$ corresponding to the rth column of $\mathbf{B}$) be ℓ_r, then we form a new set of basic variables $\mathcal{B}' := \mathcal{B} \bigcup \{j\} \setminus \{\ell_r\}$, and a corresponding new basis $\mathbf{B}' = \mathbf{A}_{\mathcal{B}'}$. That this new basis $\mathbf{B}'$ is both feasible and has cost less than or equal to the previous basis can be seen by first recalling the following inversion rule

$$(\mathbf{B} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{B}^{-1} - \frac{1}{1 + \mathbf{v}^T \mathbf{B}^{-1} \mathbf{u}} \mathbf{B}^{-1} \mathbf{u} \mathbf{v}^T \mathbf{B}^{-1}$$

and writing the new matrix in the form

$$\mathbf{B}' = \mathbf{B} + (\mathbf{a}_j - \mathbf{a}_{\ell_r}) \mathbf{e}_r^T$$

with $\mathbf{e}_r$ the rth column of the identity matrix. From this we see that

$$\begin{aligned}
(\mathbf{B}')^{-1} &= \mathbf{B}^{-1} - \frac{1}{1 + [\mathbf{B}^{-1}\mathbf{a}_j]_r - [\mathbf{B}^{-1}\mathbf{a}_{\ell_r}]_r} \mathbf{B}^{-1}(\mathbf{a}_j - \mathbf{a}_{\ell_r}) \mathbf{e}_r^T \mathbf{B}^{-1} && (21) \\
&= \mathbf{B}^{-1} - \frac{1}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} \mathbf{B}^{-1}(\mathbf{a}_j - \mathbf{a}_{\ell_r}) \mathbf{e}_r^T \mathbf{B}^{-1} && (22)
\end{aligned}$$

This gives the new solution

$$\begin{aligned}
\mathbf{x}'_{\mathcal{B}'} &= \mathbf{B}^{-1}\mathbf{b} - \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} \mathbf{B}^{-1}(\mathbf{a}_j - \mathbf{a}_{\ell_r}) && (23) \\
&= \mathbf{B}^{-1}\mathbf{b} - \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} (\mathbf{B}^{-1}\mathbf{a}_j - \mathbf{e}_r) && (24)
\end{aligned}$$

which is feasible by construction. Noting also that

$$\mathbf{c}_{\mathcal{B}'} = \mathbf{c}_{\mathcal{B}} + \mathbf{e}_r([\mathbf{c}]_j - [\mathbf{c}]_{\ell_r})$$

we see the new cost is

$$\begin{aligned}
\mathbf{c} \cdot \mathbf{x}' &= [\mathbf{c}_{\mathcal{B}} + \mathbf{e}_r([\mathbf{c}]_j - [\mathbf{c}]_{\ell_r})]^T \left(\mathbf{B}^{-1}\mathbf{b} - \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} (\mathbf{B}^{-1}\mathbf{a}_j - \mathbf{e}_r) \right) && (25) \\
&= \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1}\mathbf{b} + ([\mathbf{c}]_j - [\mathbf{c}]_{\ell_r})[\mathbf{B}^{-1}\mathbf{b}]_r - \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1}\mathbf{a}_j + \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} [\mathbf{c}]_{\ell_r} - \frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r} ([\mathbf{B}^{-1}\mathbf{a}_j]_r - 1)([\mathbf{c}]_j - [\mathbf{c}]_{\ell_r}) \\
&= \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1}\mathbf{b} + \underbrace{\frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r}}_{\geq 0} \underbrace{([\mathbf{c}]_j - \mathbf{c}_{\mathcal{B}}^T \mathbf{B}^{-1}\mathbf{a}_j)}_{<0} && (26)
\end{aligned}$$

4.7 The Simplex Algorithm (Chapter 5)

With the pivoting technique introduced in the previous chapter we are now ready to define the simplex algorithm. The algorithm starts by selecting a feasible basis (or if none exists, terminating and stating so). It then pivots until either it finds an optimal basic feasible solution, or it determines that the problem is unbounded below. The key things left yet unspecified in this description is which j among those satisfying (19) is selected (and possible selection among the multiple rs with the same minimum value in (20)).

4.8 Finding an Initial Feasible Basis

The book offers two methods for finding an initial basis. In one, the "big-M method" one augments the original problem to

$$\min\{\mathbf{c}\cdot\mathbf{x}+\gamma\mathbf{1}\cdot\mathbf{s}\,|\,\mathbf{A}\mathbf{x}+\mathbf{s}=\mathbf{b},\ \mathbf{s}\geq\mathbf{0},\mathbf{x}\geq\mathbf{0}\}$$

and chooses a sufficiently large γ. The initial feasible basis is then selected as the $\mathbf{s}$ variables. If this altered problem has an unbounded solution or an optimal solution with $s_j>0$ for some j, then the the original problem is either unbounded or has no feasible solution (provided γ was large enough). If instead the optimal solution has $\mathbf{s}=\mathbf{0}$ then the modified problem's solution is also an optimal solution to the original problem.

In the second method called the "two-phase" method, one starts with a different modified problem

$$\min\{\gamma\mathbf{1}\cdot\mathbf{s}\,|\,\mathbf{A}\mathbf{x}+\mathbf{s}=\mathbf{b},\ \mathbf{s}\geq\mathbf{0},\mathbf{x}\geq\mathbf{0}\}$$

starting with the initial basic feasible solution $\mathbf{x}=\mathbf{0},\ \mathbf{s}=\mathbf{b}$. This has an optimal solution (it's bounded below by 0). If this optimal solution is 0 then $\mathbf{1}\cdot\mathbf{s}=0$ and the optimal solution gives a feasible basis for the original problem. If the optimal solution is not 0, then there is no feasible solution for the original problem.

4.9 Cycling and Pivot Selection Rules to Avoid Cycling

Despite the fact that there are a finite number of feasible bases, and that each pivot operation between them guarantees a new cost that is $\leq$ to the old cost, cycling is still a possibility in linear programming. It occurs when a chain of pivots give the same cost, and the solution at the end of the chain $\mathbf{x}_{n+K}$ is the same as the solution at the beginning of the chain $\mathbf{x}_n$. Thus, in order to study the cycling phenomenon, it is of interest to study those pivots which leave the cost unchanged. As can be seen from (26), the only situation in which the cost doesn't decrease is when

$$\min_{r|[\mathbf{B}^{-1}\mathbf{a}_j]_r>0}\frac{[\mathbf{B}^{-1}\mathbf{b}]_r}{[\mathbf{B}^{-1}\mathbf{a}_j]_r}=0$$

Or equivalently, when the solution before the pivot operation $\mathbf{x}_{\mathcal{B}}=\mathbf{B}^{-1}\mathbf{b},\ \mathbf{x}_{\mathcal{N}\setminus\mathcal{B}}=\mathbf{0}$ had some basic variables (some $x_i,\ i\in\mathcal{B}$) which were 0, and thus was degenerate. For a cycle to occur, then, a long chain of degenerate solutions of equal cost must occur. In fact, by (24), the solution itself (which is degenerate) does not change during pivots in a cycle: different variables i with solution value $x_i=0$ are just being moved in and out of the basis.

We shall now prove if we select the column to pivot in j (among the possible candidates) according to an appropriate rule and if we choose r among the minima according to an appropriate rule we can avoid cycling.

The following simple rules work: *Choose the first j for which (19) is true (i.e. the one with the least index j)* and *Choose among those r satisfying (20) the one with the smallest ℓ_r (i.e. the smallest column index in $\mathbf{A}$)*. (If ordering is kept consistent when a subset of columns of $\mathbf{A}$ are taken into $\mathbf{B}$, then this corresponds to the smallest index r.) Using these pivot selection rules guarantee that, if the simplex algorithm is started from a feasible solution, it will converge to the optimal solution (or detect unboundedness) after a finite number of iterations.

To prove this, we assume that there would be a cycle and show that a contradiction is reached. Let $\mathcal{S}$ be the set of variables that are kept in the basis for the whole cycle (i.e. are never pivoted out), and let $\mathcal{B}_n, n\in\{1,\ldots,K\}$ be the sets of indices of basic variables at each time instant n in the first cycle, so that

$$\mathcal{S}:=\bigcap_{n=1}^{K}\mathcal{B}_n$$

Define the set of variables that cycle in and out of the basis to be

$$\mathcal{T}:=\bigcup_{n=1}^{K}\mathcal{B}_n\setminus\mathcal{S}$$

As explained above, any solution $\mathbf{x}$ during this cycle has $x_i=0$ for $i\in\mathcal{T}$. Let $q=\max\mathcal{T}$ be the largest index variable that is pivoted in. Let the basis just before the variable q is pivoted in be $\mathbf{B}_1=\mathbf{A}_{\mathcal{B}_1}$. Let the basis when q gets pivoted out be $\mathbf{B}_2=\mathbf{A}_{\mathcal{B}_2}$ and let the variable pivoted in be s. Note that if we define $\mathbf{y}_{\mathcal{B}_2}=\mathbf{B}_2^{-1}\mathbf{a}_s$ and $y_s=-1$ and $\mathbf{y}_{\mathcal{N}\setminus\mathcal{B}\setminus\{s\}}=\mathbf{0}$, then we must have $\mathbf{A}\mathbf{y}=\mathbf{0}$. Inspecting the inner product

$$(\mathbf{c}^T-\mathbf{c}_{\mathcal{B}_1}^T\mathbf{B}_1^{-1}\mathbf{A})\mathbf{y}=\mathbf{c}^T\mathbf{y}=\mathbf{c}_{\mathcal{B}_2}^T\mathbf{B}_2^{-1}\mathbf{a}_s-c_s=-[\mathbf{c}^T-\mathbf{c}_{\mathcal{B}_2}^T\mathbf{B}_2^{-1}\mathbf{A}]_s>0$$

because s had to be selected to be moved into the basis. But removing the middle parts of this chain of relations we see that we must have

$$[\mathbf{c}^T - \mathbf{c}_{\mathcal{B}_1}^T \mathbf{B}_1^{-1} \mathbf{A}]_k y_k > 0$$

for some k. Because $\mathbf{c}^T - \mathbf{c}_{\mathcal{B}_1}^T \mathbf{B}_1^{-1} \mathbf{A}$ has a zero at any element in the basis $\mathcal{B}_1$, we see that k can not be a variable in $\mathcal{B}_1$. Since tmhe only non-zero elements in $\mathbf{y}$ were with indices in $\mathcal{B}_2 \bigcup \{s\}$ we see that we must have $k \in \mathcal{B}_2 \bigcup \{s\}$. Since this implies that k moves in and out of the basis, we must have $k \in \mathcal{T}$, so that $k \leq q$. But because q leaves the basis $\mathbf{B}_2$ we must have $y_q = [\mathbf{B}_2^{-1} \mathbf{a}_s]_r(q) > 0$. Furthermore, by virtue that q enters the basis $\mathbf{B}_1$, we must have $[\mathbf{c}^T - \mathbf{c}_{\mathcal{B}_1}^T \mathbf{B}_1^{-1} \mathbf{A}]_k < 0$ which shows that $k \neq q$. Furthermore, $y_s = -1$ and since $s < q$ and q was selected to be pivoted into $\mathbf{B}_1$ we must have $[\mathbf{c}^T - \mathbf{c}_{\mathcal{B}_1}^T \mathbf{B}_1^{-1} \mathbf{A}]_s \geq 0$ so $k \neq s$ too. Now, since $k \leq q$ we must have $[\mathbf{c}^T - \mathbf{c}_{\mathcal{B}_1}^T \mathbf{B}_1^{-1} \mathbf{A}]_k > 0$ (because k could have been selected during the pivot that moved in q because it has a lower index than q), so that we must also have $y_k = [\mathbf{B}_2^{-1} \mathbf{a}_s]_{r(k)} > 0$, but this together with the fact that the solution $x_k = 0$ (since $k \in \mathcal{T}$). But these imply that k should have been selected to be pivoted out instead of q, which is a contradiction.