Extreme Points, Corners, and Basic Feasible Solutions

CS 149 Staff

October 11, 2008

1 Optimal Solutions

We have proved the equivalence of corners, extreme points and basic feasible solutions. However, there is another important part of the geometry of the feasible region that we must examine, namely **directions**. Directions, intuitively, are vectors that you can follow infinitely and stay in the feasible region. If our feasible region in $\mathbb{R}^2$ is the first quadrant, then the set of directions for this region is all $x \geq 0$. Let us define the set of directions, $C(P)$, formally:
For $P \neq \emptyset$:

$$C(P) := \{y \in \mathbb{R}^n | \forall x \in P, \lambda > 0, x + \lambda y \in P\}.$$

Remark 1. For a LP in normal form: $C(P) = \{y \in \mathbb{R}^n | Ay = 0 \text{ and } y \geq 0\}$

Proof. If we have a $y \in C(P) := \{y \in \mathbb{R}^n | \forall x \in P, \lambda > 0, x + \lambda y \in P\}$, we know that $A(x + \lambda y) = b$. We also know, since $x \in P$, that $Ax = b$. Then;
$Ax + A\lambda y = b \Rightarrow b + A\lambda y = b \Rightarrow \lambda Ay = 0 \Rightarrow Ay = 0$.
Now, the other direction of the proof is trivial; given $Ax = b$ and $y \in C(P) = \{y \in \mathbb{R}^n | Ay = 0 \text{ and } y \geq 0\}$ for any $\lambda > 0$ we have $Ax + \lambda Ay = b \Rightarrow A(x + \lambda y) = b \Rightarrow x + \lambda y \in P$.

□

Now that we have this, we can state an important theorem:

Theorem 2. *Given $P = \{x | Ax = b \wedge x \geq 0\}$ we have $P = \kappa(\epsilon(P)) + C(P)$*

The intuition is that P can be spanned by the convex hull of its extreme points together with the set of its direction vectors.

Corollary 3. *For $P = \{x | Ax = b \wedge x \geq 0\}$, there exists at least one optimal solution x, and it is an extreme point. ($P \neq \emptyset$)*

Proof. Suppose x^* is optimal for $c^T x$ in a minimization problem. Let $\epsilon(P) = \{a^1, a^2, \ldots a^r\}$ and $y \in C(P)$. Then $\exists \lambda^1, \lambda^2, \ldots \lambda^r : x = \sum \lambda^i a^i + \beta\, y$ where $\sum \lambda^i = 1$

and $\lambda^i \geq 0$. x^* requires that $c^T\ y$ is strictly positive (otherwise we could increase β and the objective would be better, hence x^* would not be optimal). So we can simply write $x^* = \sum \lambda^i a^i$. Then the optimal value $c^T x = \sum \lambda^i c^T a^i s \Rightarrow c^T a^i$ are the same $\forall i$ and $\lambda^i > 0$. $\square$

This corollary suggests an algorithm: We could enumerate all the BFS's (extreme points), compute their costs, and take the cheapest one. However, if we have m variables and n constraints, there are $\begin{pmatrix} n \\ m \end{pmatrix}$ BFS's which is exponential. Enumerating exponential number of BFS's is not tractable so we need a search algorithm that is actually guided in some way.

2 Basis Changes

2.1 Vector to another feasible solution

We are going to start with a bfs, x^0, and construct a vector, y, to another solution, x, and then show when this other solution will be another bfs.

Take a bfs x^0, spliting it into basic and non-basic parts, $x^0 = \begin{pmatrix} x_B^0 \\ x_N^0 \end{pmatrix}$ where $x_B^0 \geq 0, x_N^0 = 0$, and $A_B * x_B^0 = b$. Now let's take another feasible x: $x \in \mathbb{R}^n, Ax = b$. Consider the vector from x^0 to x: $y := x - x^0$. $Ay = A(x - x^0) = Ax - Ax^0 = b - b = 0$ and also $0 = Ay = A_B y_B + A_N y_N$. For the non-basic part of y, $y_N = x_N$ since $y_N = x_N - x_N^0$ and $x_N^0 = 0$. For the basic part $y_B = -A_B^{-1} A_N x_N$. So we can write $y = \begin{pmatrix} -A_B^{-1} A_N x_N \\ x_N \end{pmatrix}$.

x can be any feasible (not necessarily basic at this point) solution, but we would like it to be basic. In order to get this, we are going to set $x_N = e_k$, where e_k is all zeros except for a 1 in the kth position. We are going to need the corresponding non-basic column, so we define $t = N(k)$.[1]

To define our y, we still need to find out what y_B is. Remember $y_N = x_N = e_k$, we are just selecting the kth column out of A_N, which is a^t.[2] Then $y_B = -A_B^{-1} A_N y_N$ can be written as $y_B = -A_B^{-1} a^t \implies y = \begin{pmatrix} -A_B^{-1} a^t \\ e_k \end{pmatrix}$.

Now we have our y, and we know that $x = x^0 + \lambda y$ satisfies $Ax = b$. All we need for x to be feasible is to have $x \geq 0$. The x_N part will always be greater than 0, since $x_N = x_N^0 + \lambda y_N = 0 + \lambda e_k > 0$. Then the question is when is $x_B = x_B + \lambda y_B$ greater than zero? Let us consider this on an element-by-element basis. Trivially, any element of y that is greater than zero will never give us a problem, since the

[1] The k'th column in the non-basis.

[2] The t'th column of the matrix A.

corresponding x_B will be the sum of 2 positive numbers. Zero elements also cause no problems, so its only the negative elements that interest us.

We want to find a λ such that one element of $x_B = 0$. This element is in the basis,[3] since we are talking about x_B, so let's call it element $B(j)$. We have $0 = x_{B(j)} = x^0_{B(j)} + \lambda y_{B(j)} \Rightarrow \lambda = -\frac{x^0_{B(j)}}{y_{B(j)}}$. We will have $x \geq 0$ if $\lambda \leq -\frac{x^0_{B(j)}}{y_{B(j)}} \forall j$ such that $y_{B(j)} < 0$ and we have that at least one formerly basic component of x^0 that is not zero in x. Thus, our range of acceptable λ is $0 \leq \lambda \leq min\{-\frac{x^0_{B(j)}}{y_{B(j)}} | y_{B(j)} < 0\}$.

2.2 Adjacent BFS

Now we have a feasible $x = x^0 + \lambda y$ for some range of λ. Given this, we can state and prove an important theorem:

Theorem 4. *If $\exists y_{B(j)} < 0$, we choose λ as large as possible ($\lambda < \infty$). Then, $x^1 = x^0 + \lambda y$ is a basic feasible solution and the corresponding basis is given by:*

$B^*(i) := B(i)$ *if* i $\neq$ r
$B^*(r) := t = N(k)$
r is chosen such that $\lambda = \frac{x^0_{B(r)}}{-y_{B(r)}} = min\{\frac{x^0_{B(i)}}{-y_{B(j)}} | y_{B(j)} < 0\}$.
If a new solution x^1 is obtained from x^0, we say the two corners are adjacent.

Proof. First, let us examine the non-basic[4] elements.
Define $N^*(i) = \left\{ \begin{array}{ll} N(i) & i \neq k \\ B(r) & i = k \end{array} \right\}$
For all $i \neq k$, we know that $x^1_{N^*(i)} = x^0_{N(i)} + \lambda e_{k\ N(i)} = 0 + \lambda 0$.
If $i = k$, we know that $x^1_{N^*(i)} = x^0_{B(r)} + \lambda y_{B(r)} = x^0_{B(r)} + \frac{x^0_{B(r)}}{-y_{B(r)}} y_{B(r)}$ from our definition of λ. Then, $x^1_{N^*(i)} = x^0_{B(r)} - x^0_{B(r)} = 0$.
So, all of the non-basic elements are zero. We have our injective functions defining the basis and the non-basis, and we have $x^1_{N^*} = 0$. Therefore we have a bfs. $\square$

Now, we provide an example that starts with bfs x^0, calculates λ, defines direction vector y and moves to an adjacent bfs solution x^1.

Below is the sample tableau:

1	0	0	1	1	3
0	1	0	−1	1	1
0	0	1	1	−1	1

[3]Still the basis of x_0.
[4]In x^1's basis, as defined above.

This tableau corresponds to the bfs $x^0 = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

We choose the second non-basic column since it has a negative coefficient in the objective and scan the positive entries. The first 2 are positive, and we choose the one that minimizes the ratio $\frac{b_i}{a_i^t}$ (also remember that $x_B^0 = b$). Our two candidates are 3 and 1, so we take $1 = \lambda$ and pivot around the highlighted element:

$$\begin{array}{ccccc|c}
\hline
1 & 0 & 0 & 1 & 1 & 3 \\
0 & 1 & 0 & -1 & \mathbf{1} & 1 \\
0 & 0 & 1 & 1 & -1 & 1
\end{array}$$

We can read out our y vector:$y = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

Notice that the basic elements of y are the negatives of the corresponding elements in $-y_B$, and that since we have chosen the second non-basic column, the first non-basic entry is zero and the second is one.
Lets do the steps out:

$$\begin{array}{ccccc|c}
\hline
1 & 0 & 0 & 1 & 1 & 3 \\
0 & 1 & 0 & -1 & \mathbf{1} & 1 \\
0 & 1 & 1 & 0 & 0 & 2
\end{array}$$

$$\begin{array}{ccccc|c}
\hline
1 & -1 & 0 & 2 & 0 & 2 \\
0 & 1 & 0 & -1 & \mathbf{1} & 1 \\
0 & 1 & 1 & 0 & 0 & 2
\end{array}$$

We can now read out our new solution:$x^1 = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$, and verify that:

$$x^1 = x^0 + \lambda y = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$