Basic solutions

Consider the canonical LP

$$\max \boldsymbol{c}\cdot\boldsymbol{x}, \text{ s.t. } A\boldsymbol{x}=\boldsymbol{b}, \;\; \boldsymbol{x}\in\mathbb{R}^n_+, \, \boldsymbol{b}\in\mathbb{R}^m_+, \, n\geq m.$$

Assume that the rows of the $m\times n$ matrix A are linearly independent, for otherwise the system of equations $A\boldsymbol{x}=\boldsymbol{b}$ is either redundant, that is the number of rows can be reduced, or it is inconsistent, i.e. the problem is unfeasible.

Suppose, $\boldsymbol{x}$ is a feasible solution. Then if $\boldsymbol{a}^j$, $j=1,\ldots,n$ are the *columns* of A, i.e. vectors in $\mathbb{R}^m$, the fact that $A\boldsymbol{x}=\boldsymbol{b}$ means

$$x_1\boldsymbol{a}^1+x_2\boldsymbol{a}^2+\ldots+x_n\boldsymbol{a}^n=\boldsymbol{b},$$

i.e. $\boldsymbol{b}$ is a linear combination of the columns of A with non-negative coefficients x_j. Equivalently, $\boldsymbol{b}$ is spanned by these columns, with the array of non-negative coefficients $\boldsymbol{x}$.

Some components x_j may be zero and therefore $\boldsymbol{a}^j$ does not really belong to the spanning set of vectors. So, the set of those $j's$, a subset of $\{1,\ldots,n\}$, where $x_j>0$ is called *basis*, the corresponding components $x_j>0$ – *basic* components, and the corresponding columns $\boldsymbol{a}^j$ of A – *basic* columns. This is all relative to the feasible solution $\boldsymbol{x}$.

> Definition: A feasible solution solution $\boldsymbol{x}$ is called *basic* if either $\boldsymbol{x}=\boldsymbol{0}$, or the columns of A, corresponding to *nonzero* components of $\boldsymbol{x}$ in the above linear combination are *linearly independent.*

It follows that a basic solution cannot have more than m nonzero components. Indeed, columns of A are vectors in $\mathbb{R}^m$. The dimension of space *is* the maximum number of linearly independent vectors the space can host. So, in general, a basic solution will have no more than m positive components, and as a very crude estimate, there can never be more than $\sum_{k=0}^m \begin{pmatrix} n \\ k \end{pmatrix}$ basic solutions. (We consider the cases when a basic feasible solution has $k=0,1,\ldots,m$ positive components and sum over them.)

Of course, $\boldsymbol{x}=0$ only if $\boldsymbol{b}=0$. Otherwise, intuitively, for a "typical" $\boldsymbol{b}$, a "typical" basic solution would haves exactly m positive components.

The next theorem is of key importance: it tells one that in order to seek feasible or optimal solutions of an LP, it suffices to confine oneself to basic ones (whose number is finite) only. Then the way of solving LPs would be simply a clever inspection of one basic solution after another; this is exactly what the simplex method does.

> Theorem on basic solutions: (i) If the problem is feasible, there exists a basic feasible solution (BFS). (ii) If the problem is optimizable (has optimal solution), there exists a basic optimal solution (BOS).

Proof of (i): Of all the feasible solutions, take one with the *smallest possible* number of nonzero (positive) components: such a solution $\boldsymbol{x}$ always exists. Then if $\boldsymbol{x}=\boldsymbol{0}$, it is by definition basic, and there's nothing left to prove. So suppose, $\boldsymbol{x}\neq\boldsymbol{0}$,, that is

$$x_\alpha\boldsymbol{a}^\alpha+x_\beta\boldsymbol{a}^\beta+\ldots=\boldsymbol{b},$$

for some columns $\boldsymbol{a}^\alpha,\boldsymbol{a}^\beta,\ldots$ of A.

Now the claim is that either the columns $\boldsymbol{a}^\alpha,\boldsymbol{a}^\beta,\ldots$ are linearly independent (which makes $\boldsymbol{x}$ a BFS), or the total number of nonzero components of x can be reduced, which contradicts the choice of $\boldsymbol{x}$ as a feasible solution, having the smallest number of nonzero components.

Indeed, suppose

$$\lambda_\alpha \boldsymbol{a}^\alpha + \lambda_\beta \boldsymbol{a}^\beta + \ldots = 0,$$

where at least one of the λs, say λ_α is positive. (This can always be achieved by multiplying the last equation by -1 if necessary. In fact, the equation can be multiplied by any real number and will still retain zero in the right-hand side, so the array of λ's is defined up to a real multiple). So, having λ's fixed, with without loss of generality $\lambda_\alpha > 0$, we can find a small positive number θ, such that

$$(x_\alpha - \theta\lambda_\alpha)\boldsymbol{a}^\alpha + (x_\beta - \theta\lambda_\beta)\boldsymbol{a}_\beta + \ldots = \boldsymbol{b},$$

- which results from subtraction of the last quoted equation multiplied by θ from the penultimate one - is another feasible solution $\boldsymbol{x}_\theta$, i.e al the coefficients in brackets are still non-negative. This is one step away from getting a contradiction: we shall now increase θ until one (or more) of the coefficients in brackets become zero. As soon as it happens we stop. This provides a feasible solution that has fewer positive components than $\boldsymbol{x}$ – contradiction.

To express the above argument rigorously, the array of λ's can be extended to all $j = 1, \ldots, n$ by making extra assignments $\lambda_j = 0$ (whenever $x_j = 0$). Letting then $\boldsymbol{\lambda} = (\lambda_1, \ldots \lambda_n)$, we have $\boldsymbol{x}_\theta = \boldsymbol{x} - \theta\boldsymbol{\lambda}$.

Clearly, choosing

$$\theta = \min_{j=1,\ldots,n:\, x_j, \lambda_j > 0} \frac{x_j}{\lambda_j}$$

does the job. Namely, $\boldsymbol{x}_\theta$ is feasible and has one nonzero component less than $\boldsymbol{x}$, which is the contradiction to how $\boldsymbol{x}$ has been chosen. And what has led to the contradiction was the assumption that the basic columns of A, defined relative to $\boldsymbol{x}$ were linearly dependent. So they are not, i.e. $\boldsymbol{x}$ is a basic feasible solutions.

Proof of (ii): Goes in the same way. Of all the optimal solutions, take any with the *smallest possible* number of nonzero components, call it $\boldsymbol{x}$. If $\boldsymbol{x} = \boldsymbol{0}$, it's basic. Otherwise, suppose it has non-zero components $x_{\alpha,\beta,\ldots}$, i.e

$$x_\alpha \boldsymbol{a}^\alpha + x_\beta \boldsymbol{a}^\beta + \ldots = \boldsymbol{b}.$$

If the columns $\boldsymbol{a}^{\alpha,\beta,\cdots}$ are linearly independent, $\boldsymbol{x}$ is a BOS. If they are not, then

$$\lambda_\alpha \boldsymbol{a}^\alpha + \lambda_\beta \boldsymbol{a}^\beta + \ldots = 0.$$

Once again, consider an n-vector $\boldsymbol{\lambda}$ by augmenting the array $\lambda_\alpha \boldsymbol{a}^\alpha, \lambda_\beta, \ldots$ by defining $\lambda_j = 0$ if $x_j = 0$. a

Now, for $\theta \in \mathbb{R}$, with a small enough *absolute value*, the solution $\boldsymbol{x}_\theta = \boldsymbol{x} - \theta\boldsymbol{\lambda}$ is still feasible.

Looking at the value $V(\boldsymbol{x}_\theta) = \boldsymbol{c} \cdot \boldsymbol{x}_\theta = V(\boldsymbol{x}) + \theta\boldsymbol{c} \cdot \boldsymbol{\lambda}$, one concludes that it has to be $\boldsymbol{c} \cdot \boldsymbol{\lambda} = 0$. Indeed, otherwise $\boldsymbol{x}$ cannot be optimal: by choosing the sign of θ (plus or minus) we could achieve both $V(\boldsymbol{x}_\theta) > V(\boldsymbol{x})$ and $V(\boldsymbol{x}_\theta) < V(\boldsymbol{x})$.

Hence, we conclude that $V(\boldsymbol{x}_\theta) = V(\boldsymbol{x})$, i.e. as long as $\boldsymbol{x}_\theta$ is feasible, it is optimal as well.

If so, we repeat precisely the same trick as in part (i). One can choose θ such that the number of positive components of the optimal $\boldsymbol{x}_\theta$ becomes less than that of $\boldsymbol{x}$, by choosing θ equal to the *minimum positive ratio* x_j/λ_j (needless to say, over those $\lambda_j > 0$).

This contradicts the choice of $\boldsymbol{x}$, so the columns $\boldsymbol{a}^{\alpha,\beta,\cdots}$ have to be linearly independent, which means $\boldsymbol{x}$ is a BOS.■