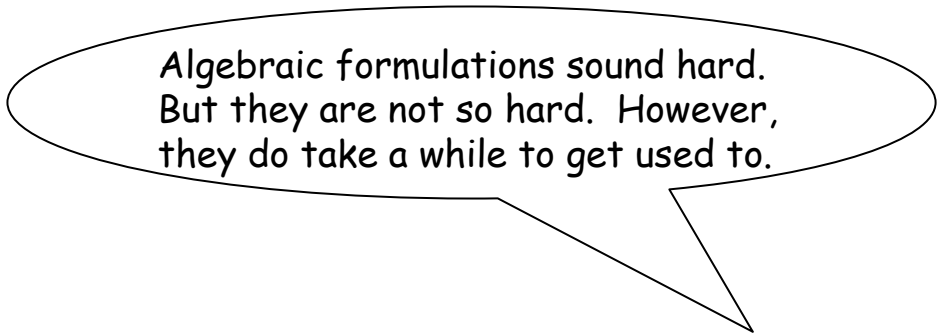
Algebraic Formulations

Usually in class, we describe linear programs by writing them out fully. This is fine for small linear programs, but it doesn't work when the linear programs are very large. In that case, it helps to use algebraic formulations.



Cleaver, an MIT Beaver

On creating algebraic formulations

- **When we create algebraic formulations, we rely on substituting notation for some of the coefficients. Let's start with an example of a linear program.**

Minimize

$$500\,x_1 + 200\,x_2 + 250\,x_3 + 125\,x_4$$

subject to

$$50{,}000\,x_1 + 25{,}000\,x_2 + 20{,}000\,x_3 + 15{,}000\,x_4 \geq 1{,}500{,}000$$

$$0 \leq x_1 \leq 20$$

$$0 \leq x_2 \leq 15$$

$$0 \leq x_3 \leq 10$$

$$0 \leq x_4 \leq 15$$

This is the MSR example from lecture 1.

Cleaver

More on the MSR Problem

- **Need to choose ads to reach at least 1.5 million people**
  - **Minimize Cost**
  - **Upper bound on number of ads of each type**

	TV	Radio	Mail	Newspaper
Audience Size	**50,000**	**25,000**	**20,000**	**15,000**
Cost/Impression	**$500**	**$200**	**$250**	**$125**
Max # of ads	**20**	**15**	**10**	**15**

In the MSR problem, we wanted to determine the number of ads of four different types of advertising media.

We let x_1, x_2, x_3, and x_4 denote the number of ads on TV, Radio, Mail, and Newspaper.

The LP Formulation again

Minimize

$$500\,x_1 + 200\,x_2 + 250\,x_3 + 125\,x_4$$

subject to

$$50{,}000\,x_1 + 25{,}000\,x_2 + 20{,}000\,x_3 + 15{,}000\,x_4 \geq 1{,}500{,}000$$

$$0 \leq x_1 \leq 20$$

$$0 \leq x_2 \leq 15$$

$$0 \leq x_3 \leq 10$$

$$0 \leq x_4 \leq 15$$

The objective is the cost of advertising. The first constraint says that the number of people who see the ads is at least 1.5 million.

The remaining four constraints give upper and lower bounds on the number of showings of each of the four ads.

Transforming into an algebraic problem

- **We'll transform this problem into an algebraic version in a couple of stages. Then we'll show how to do it all at once.**

- **So, let's start with the four upper bound constraints. Suppose that we let d = (d_1, d_2, d_3, d_4) = (20, 15, 10, 15). We can then write the linear program as follows:**

Minimize

$$500\, x_1 + 200\, x_2 + 250\, x_3 + 125\, x_4$$

subject to

$$50{,}000\, x_1 + 25{,}000\, x_2 + 20{,}000\, x_3 + 15{,}000\, x_4 \geq 1{,}500{,}000$$

$$0 \leq x_j \leq d_j \text{ for } j = 1 \text{ to } 4.$$

It looks like d_j is a variable, but it isn't. It's called a "parameter" and it means that there is an associated value stored for it somewhere.

Ollie,
the computationally
wise owl.

Parameters versus decision variables

I don't get it. The d's don't look like numbers to me.

Most students (and others) find this confusing at the beginning. But after a while, one gets used to it.

For students who have seen linear algebra, it's pretty similar to when one first sees a system of linear equations expressed as $Ax = b$.

Tim, the turkey

Ollie, the computationally wise owl.

More on the algebraic formulation

Minimize $$500\,x_1 + 200\,x_2 + 250\,x_3 + 125\,x_4$$

subject to $$50{,}000\,x_1 + 25{,}000\,x_2 + 20{,}000\,x_3 + 15{,}000\,x_4 \geq 1{,}500{,}000$$

$$0 \leq x_j \leq d_j \text{ for } j = 1 \text{ to } 4.$$

Another advantage of the algebraic formulation is that the formulation becomes "independent" of the data. For example, if we were to change the upper bounds on the x's, this more algebraic version would still be valid.

Actually, it won't be the algebraic version until we get rid of almost all of the numbers. We will permit the number 0 at times, plus numbers for the indices.

Making the remaining constraint more algebraic

Minimize $500\,x_1 + 200\,x_2 + 250\,x_3 + 125\,x_4$

subject to $a_1\,x_1 + a_2\,x_2 + a_3\,x_3 + a_4\,x_4 \geq b$

$0 \leq x_j \leq d_j$ for $j = 1$ to 4.

Let a_j be the coefficient of x_j in the constraint. And let b denote the RHS of the constraint. We then can rewrite the constraint.

It doesn't look simpler than the old version. But it would if there were 1000 variables instead of just 4.

Ollie

Nooz

Transforming the cost coefficients

Minimize $c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4$

subject to $a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 \geq b$

$0 \leq x_j \leq d_j$ **for j = 1 to 4.**

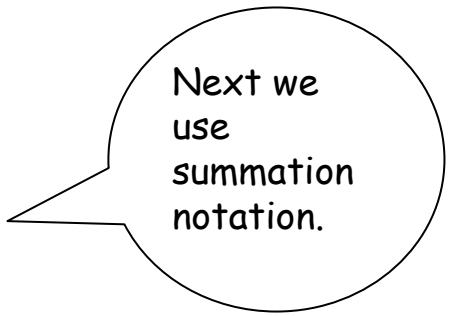
Let c_j be the cost coefficient of x_j We now rewrite the objective.

Using Summation Notation

Minimize $\sum_{j=1}^{4} c_j x_j$

subject to $\sum_{j=1}^{4} a_j x_j \geq b$

$0 \leq x_j \leq d_j$ for j = 1 to 4.



Replacing the number of variables.

Minimize $\sum_{j=1}^{n} c_j x_j$

subject to $\sum_{j=1}^{n} a_j x_j \geq b$

$$0 \leq x_j \leq d_j \quad \text{for } j = 1 \text{ to } n.$$

Finally, we use n to represent the number of variables.

Yes. I know it looks much more abstract that the original formulation. But the abstraction means that this formulation is correct for many different possible choices of the data.

Ollie

Nooz

Summary of the transformation

- **Let x_j be the number of ads purchased of type j for j = 1 to n.**
- **Let a_j be the number of persons who view one ad of type j**
- **Let b be the required number of viewers to see the ads. (That is, the total number of viewers must be at least b)**
- **Let d_j be an upper bound on the number of ads purchased of type j.**

Minimize $\sum_{j=1}^{n} c_j x_j$

subject to $\sum_{j=1}^{n} a_j x_j \geq b$

$0 \leq x_j \leq d_j$ for j = 1 to n.

Minimize $500\, x_1 + 200\, x_2 + 250\, x_3 + 125\, x_4$

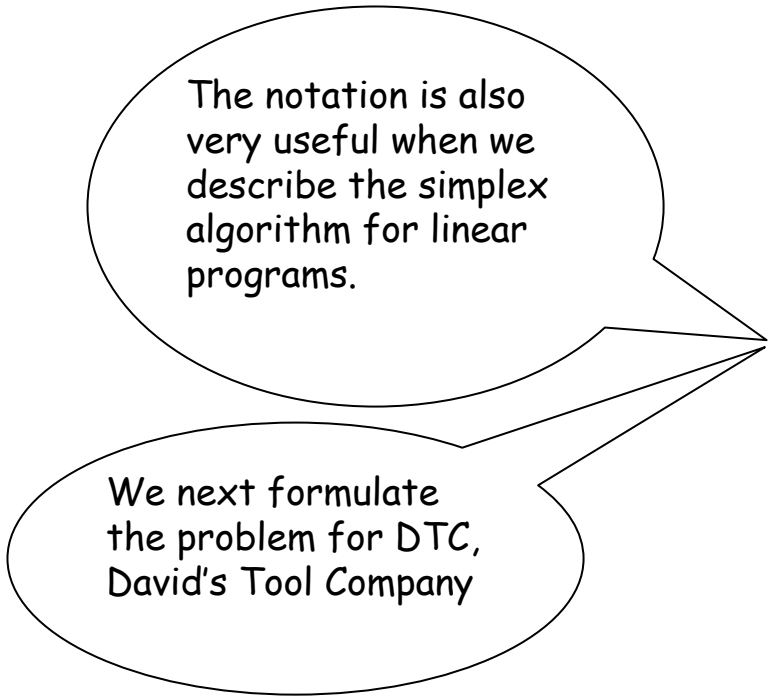
subject to $50{,}000\, x_1 + 25{,}000\, x_2 + 20{,}000\, x_3 + 15{,}000\, x_4 \geq 1{,}500{,}000$

$0 \leq x_1 \leq 20;$ $0 \leq x_2 \leq 15;$ $0 \leq x_3 \leq 10;$ $0 \leq x_4 \leq 15;$

On the reason for algebraic formulations

- **Remember that the advantage of algebraic formulations is in their ability to describe very large problems in a very compact manner. This is critical if one is to model large problems, involving thousands or perhaps millions of variables.**

- **For small problems, it seems unnecessarily cumbersome and difficult. And it is unnecessarily abstract for describing small problems.**



Ollie

Formulation of the DTC Problem (David's Tool Company)

We will write the linear program as if up the constraints are broken into two parts, the demand constraints and the resource constraints.

Maximize $z = 3K + 5S$

subject to

$2K + 3S \leq 100$	**Gathering time:**	Resource Constraints
$K + 2S \leq 60$	**Smoothing time:**	
$K + S \leq 50$	**Delivery time:**	

$K \leq 40$	**Slingshot demand:**	Bounds on variables.
$S \leq 30$	**Shield demand:**	
$K, S \geq 0$	**Non-negativity:**	

The algebraic formulation

Maximize $z = 3x_1 + 5x_2$

subject to

$2x_1 + 3x_2 \le 100$	**Gathering time:**	Resource Constraints
$x_1 + 2x_2 \le 60$	**Smoothing time:**	
$x_1 + x_2 \le 50$	**Delivery time:**	

$x_1 \le 40$	**Slingshot demand:**	Bounds on variables.
$x_2 \le 30$	**Shield demand:**	
$x_1, x_2 \ge 0$	**Non-negativity:**	

Let x_j amount of the j-th item that is produced. In the above formulation we have replaced K by x_1 and S by x_2. This will make it more easily described using algebraic notation.

The algebraic formulation

Maximize $z = \sum_{j=1}^{n} p_j x_j$

subject to

$\sum_{j=1}^{n} a_{ij} x_j \geq b_i$ for i = 1 to m.	**Gathering time:** **Smoothing time:** **Delivery time:**	Resource Constraints

$\mathbf{0 \leq x_j \leq d_j}$ **for j = 1 to n.**	**Slingshot demand:** **Shield demand:** **Non-negativity:**	Bounds on variables.

Let d_j be an upper bound on the demand for item j.
Let n denote the number of items.

Let a_{ij} be the amount of resource i used up by one unit of item j.
Let m denote the number of different resources.

Let p_j be the profit from making one unit of item j

Notation for linear programs in standard form

- **Finally, we show some conventions that are used in describing a linear programming in standard form. The conventions are used in 15.053.**

- **There are usually n variables and m equality constraints**
- **The variables are usually $x_1, \ldots, x_n$.**
- **The cost coefficients are usually $c_1, \ldots, c_n$. (Objective function coefficients are often called cost coefficients even if one is maximizing profit. Every knows that this is a weird convention, but it is commonly done.)**
- **The coefficient for x_j in constraint i is a_{ij}. The RHS is b_i.**
- **Then the LP in standard form can be written as follows:**

$$\text{maximize} \quad z = \sum_{j=1}^{n} c_j x_j$$

$$\text{subject to} \quad \sum_{j=1}^{x} a_{ij} x_j = b_i \qquad \text{for } i = 1 \text{ to } m$$

$$x_j \geq 0 \qquad \text{for } j = 1 \text{ to } n$$

Last slide

In case you were wondering, there are different ways of writing algebraic formulations. You can choose notation differently, and you can combine groups of constraints differently.

You will have a chance to practice algebraic formulations on the homework sets.

And that's the end of this tutorial. I hope it was of value to you. Bye!

Ollie

Cleaver