CS 524: Introduction to Optimization
Lecture 19 : Indicator Variables

Michael Ferris

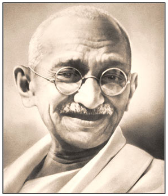
Computer Sciences Department
University of Wisconsin-Madison

October 18, 2023

Review

- In the last few lectures we have seen how to read in and write out values at compile and execution time to both 'gdx' files (e.g. $gdxin and `execute_unloaddi`) and to other format files using $ondelim, put and Jupyter notebooks (also GAMS/Connect).
- Also mentioned new variable types (e.g. `integer`, `binary`) that lead to MIP problems and outlined a way to solve such problems. Could for example update `topbrass` examples to now use integer variables for the number of tropies used.
- The next few lectures will detail more sophisticated uses of discrete variables to model new features. Today - indicator variables.

Gandhi Clothes?



- Gandhi clothes makes a set of items

$$I = \{\text{shirts, shorts, and pants}\}.$$

- A different machine must be rented (at a cost f_i) to make each of these three items.
  - Shirt machine costs \$200/month
  - Shorts machine costs \$150/month
  - Pants machine costs \$100/month.
- Gandhi can choose not to rent a machine, but then he cannot make any of that item

Gandhi Clothes

- Shirt: requires 3 hours of labor and 4 square yards of cloth; Sells for \$12 and costs \$6 to make
- Pants: requires 6 hours of labor and 4 square yards of cloth; Sells for \$15 and costs \$8 to make
- Shorts: require 2 hours of labor and 3 square yards of cloth. Sells for \$8 and costs \$4 to make
- The total amount of labor available is 150 hours, and there are 160 square yards of cloth.

Help Gandhi Out

Determine which item(s) should be manufactured, and how many of each.

Sets and Parameters

Sets

- I: Items
- R: Resources

Parameters

- r_i: Revenue of $i \in I$
- f_i: Fixed cost of $i \in I$
- c_i: Variable cost of $i \in I$
- a_{ri}: Amount of resource r required to make item i
- b_r: Total amount of resource $r \in R$

Gandhi

Variables

- x_i: Number of item $i \in I$ to make

Constraints

- Resource availability:

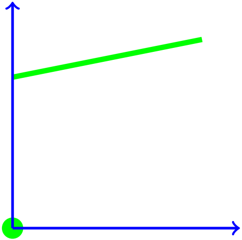
$$\sum_{i \in I} a_{ri} x_i \leq b_r \qquad \forall r \in R$$

Objective

- Maximize revenue - costs

$$\sum_{i \in I} r_i x_i - \sum_{i \in I} F_i(x)$$

Fixed Costs



$$F_i(x) = \begin{cases} 0 & x = 0 \\ f_i + c_i x & x > 0 \end{cases}$$

The Question to Ask!

- What is the shape of -$F_i(x)$?

- Since maximizing convex (minimizing concave) functions is hard, we would *not* expect there to be a trick to model this is a linear fashion
- We need to resort to integer variables!

Indicator Variables

- Idea: Use a 0-1 (binary) variable z_i to indicate if $x_i > 0$

$$\max \sum_{i \in I} r_i x_i - \sum_{i \in I} c_i x_i - \sum_{i \in I} f_i z_i$$

$$\begin{aligned}
\sum_{i \in I} a_{ri} x_i &\leq b_r && \forall r \in R \\
x_i > 0 &\Rightarrow z_i = 1 && \forall i \in I \\
x_i &\geq 0 && \forall i \in I \\
z_i &\in \{0, 1\} && \forall i \in I
\end{aligned}$$

NO!

- Logical statements like $x_i > 0 \Rightarrow z_i = 1$ are not allowed in algebraic models
- We must convert $x_i > 0 \Rightarrow z_i = 1$ to algebra

Key observation: turn on indicator

- Suppose that X, is the "set of points of interest" (whole space, integer grid, polytope, etc) and $f(x) \leq M$ for all $x \in X$
- For an $x \in X$, consider $f(x) > 0 \Rightarrow \delta = 1$, where $\delta \in \{0, 1\}$
- When $x \in X$, this is implied by the inequality:

$$f(x) \leq M\delta$$

- Since $\delta \in \{0, 1\}$, if $f(x) > 0$ the inequality can only be satisfied if $\delta = 1$ and does not change the problem since $f(x) \leq M$. If $f(x) \leq 0$, then the inequality is always satisfied regardless of the value of δ.
- In the case of interest $X = [0, M_i]$, $f(x) = x_i$ and $\delta = z_i$ resulting in the following linear inequality:

$$x_i \leq M_i z_i$$

New Gandhi

$$\max \sum_{i \in I} r_i x_i - \sum_{i \in I} c_i x_i - \sum_{i \in I} f_i z_i$$

$$\begin{aligned}
\sum_{i \in I} a_{ri} x_i &\leq b_r \quad \forall r \in R \\
x_i &\leq M_i z_i \quad \forall i \in I \\
x_i &\geq 0 \quad \forall i \in I \\
z_i &\in \{0, 1\} \quad \forall i \in I
\end{aligned}$$

What about the M's?

- $M_i = 10^4 \quad \forall i$?
- Can we make M_i smaller?
  - Smaller M_i are good!
- It gives a tighter relaxation.

Small M's Good. Big M's Baaaaaaaaaad!

- Let's look at the geometry of a simple problem

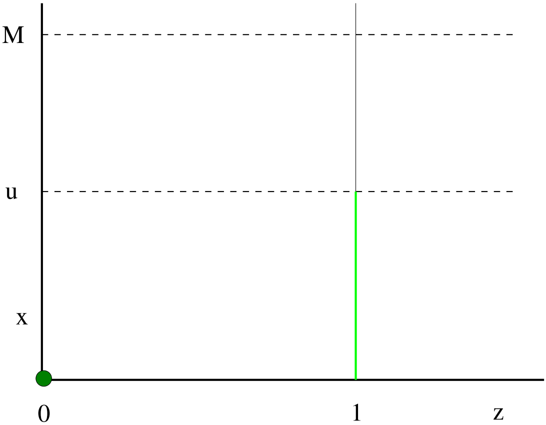
$$P = \{(z, x) : z \in \{0, 1\}, x \in \mathbb{R}_+, x \leq Mz, x \leq u\}$$

$$LP(P) = \{(z, x) : z \in [0, 1], x \in \mathbb{R}_+, x \leq Mz, x \leq u\}$$

$$\text{conv}(P) = \{(z, x) : z \in [0, 1], x \in \mathbb{R}_+, x \leq uz\}$$

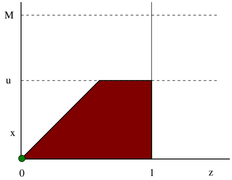
- $LP(P)$ is the standard relaxation of P
- conv(P) is the convex hull of P: the set of all convex combinations of points in P

P

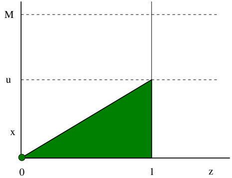


$$P = \{(z, x) : z \in \{0, 1\}, x \in \mathbb{R}_+, x \leq Mz, x \leq u\}$$

LP Versus Conv



$$LP(P) = \{(z,x) : z \in [0,1], x \in \mathbb{R}_+, x \le Mz, x \le u\}$$



$$\text{conv}(P) = \{(z,x) : z \in [0,1], x \in \mathbb{R}_+, x \le uz\}$$

- KEY: If $M = u$, $LP(P) = \text{conv}(P)$
- Extreme points of conv(P) are points in P
- Small M's good. Big M's baaaaaaaad.