Prof. Friedrich Eisenbrand
Martin Niemeier

Due Date: March 02, 2009
Discussions: March 03, 2009

Introduction to Discrete Optimization

Spring 2009

Solutions 2

Exercise 1

A factory produces two different products. To create one unit of product 1, it needs one unit of raw material A and one unit of raw material B. To create one unit of product 2, it needs one units of raw material B and two units of raw material C. Raw material B needs preprocessing before it can be used, which takes one minute per unit. At most 20 hours of time is available per day for the preprocessing. Raw materials of capacity at most 1200 can be delivered to the factory per day. One unit of raw material A, B and C has size 4, 3 and 2 respectively.

At most 130 units of the first and 100 units of the second product can be sold per day. The first product sells for 6 CHF per unit and the second one for 9 CHF per unit.

1. Formulate the problem of maximizing turnover as a linear program and solve it using Zimpl and an LP solver of your choice.
2. Can you formulate the program as a linear program with two variables as well?

Solution

1. We introduce 5 variables. Variables x and y model the amount of units of product one and two respectively that should be created. Variables a, b and c model the amount of raw material A, B and C needed.

   One unit of product one sells for 6 CHF and one unit of product two sells for 9 CHF per unit. Thus the turnover is $6x+9y$. This is our objective function.

   To create one unit of product one we need one unit of A. Thus $x \leq a$ has to hold for each feasible solution. For each unit of product one and two we need one unit of B. Thus $x + y \leq b$ must hold. Similary we get the constraint $2y \leq b$.

   Raw material B needs preprocessing that takes 1 minute per unit. At most 20 hours are available for preprocessing. This limits the maximum amount of b: $b \leq 60 \cdot 20 = 1200$.

   The total size of raw materials we can have per day is limited by 1200. One unit of raw material A, B and C has size 4, 3 and 2 respectively. This gives the constraint $4a+3b+2c \leq 1200$.

   At most 130 units of product 1 and 100 units of product two can be sold per day, thus $x \leq 130$ and $y \leq 100$ must hold. Finally all variables must be nonnegative since we cannot produce a negative amount of products or have a negative stock of raw material: $x, y, a, b, c \geq 0$.

This yields the following linear program:

$$\begin{aligned} \max \quad & 6x+9y & & \qquad (1) \\ \text{subject to} \quad & x & \le a \\ & x+y & \le b \\ & 2y & \le c \\ & b & \le 1200 \\ & 4a+3b+2c & \le 1200 \\ & x & \le 130 \\ & y & \le 100 \\ & x,y,a,b,c \ge 0 \end{aligned}$$

A ZIMPL model looks as follows:

```
# The variables of the linear program
var a >= 0 <= infinity;
var b >= 0 <= 1200;
var c >= 0 <= infinity;
var x >= 0 <= 130;
var y >= 0 <= 100;

# The objective function
maximize turnover: 6*x+9*y;

# The constraints
subto c1: x <= a;
subto c2: x+y <= b;
subto c3: 2*y <= c;
subto c4: 4*a+3*b+2*c <= 1200;
```

An optimal solution can be computed with an LP solver:

$$x = 71.4286,\ y = 100,\ a = 71.4286,\ b = 171.429,\ c = 200.$$

The value of this solution is 1328.57.

2. Consider an optimal solution x^*, y^*, a^*, b^*, c^* of the linear program (1). Observe that if we decrease a^*, b^* and c^* such that the first three constraints are satisfies with equality, the solution stays feasible and the objective value does not change. Thus this modified solution is optimal as well.

   This justifies that we consider the modified LP instead where the fist three inequalities are replaced by equalites, i.e.

$$\begin{array}{lrl} \max & 6x+9y & \\ \text{subject to} & x & = a \\ & x+y & = b \\ & 2y & = c \\ & b & \leq 1200 \\ & 4a+3b+2c & \leq 1200 \\ & x & \leq 130 \\ & y & \leq 100 \\ & x,y,a,b,c & \geq 0 \end{array}$$

By substituting a, b and c in the remaining inequalities we get a linear program with 2 variabes

$$\begin{array}{lcl} \max & 6x+9y & \\ \text{subject to} & 2y & \leq 1200 \\ & 4x+3(x+y)+2\cdot 2y & \leq 1200 \\ & x & \leq 130 \\ & y & \leq 100 \\ & x,y & \geq 0 \end{array}$$

which we can simplify to

$$\begin{array}{lcl} \max & 6x+9y & \\ \text{subject to} & 2y & \leq 1200 \\ & 7x+7y & \leq 1200 \\ & x & \leq 130 \\ & y & \leq 100 \\ & x,y & \geq 0 \end{array}$$

Exercise 2
Show that the set of feasible solutions of every linear program is convex.

Solution
As seen in the lecture, every linear program can be written in the form

$$\min\{c^T x \,:\, Ax \leq b\} \tag{2}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$.

To prove that the set of feasible solutions is convex, we need to show that for each two feasible solutions x' and x'' , every point on the line segment defined by x' and x'' is a feasible solution as well, i.e. for each $\lambda \in [0,1]$, the vector $\lambda x' + (1-\lambda)x''$ is feasible.

Let x' and x'' be feasible solutions to (2), i.e. $Ax' \leq b$ and $Ax'' \leq b$ holds. Let $\lambda \in [0,1]$.

We have

$$A(\lambda x' + (1-\lambda)x'') = \lambda Ax' + (1-\lambda)Ax'' \leq \lambda b + (1-\lambda)b = b,$$

where we used the linearity of matrix multiplication and the assumption that x' and x'' are feasible. This shows that $\lambda x' + (1-\lambda)x''$ is feasible and thus is contained in the set of feasible solutions.

Hence the set of feasible solutions of a linear program is convex.

Exercise 3
Provide an example of a convex and closed set $K \subseteq \mathbb{R}^2$ and a linear objective function $c^T x$ such that $\inf\{c^T x \;:\; x \in K\} > -\infty$ but there does not exists an $x^* \in K$ with $c^T x^* \leq c^T x$ for all $x \in K$.

Solution
Consider the set

$$K := \left\{ (a,b)^T \in \mathbb{R}^2 \;:\; a \geq 1 \;\wedge\; b \geq \frac{1}{a} \right\}.$$

The set is clearly closed. To prove that it is convex consider two points $(a,b)^T$, $(a',b')^T \in K$. Thus $b \geq \frac{1}{a}$ and $b' \geq \frac{1}{a'}$.

We have to show that for each $\lambda \in [0,1]$, the vector

$$\lambda(a,b)^T + (1-\lambda)(a',b')^T = (\lambda a + (1-\lambda)a', \lambda b + (1-\lambda)b')^T$$

is contained in K as well. Let $\tilde{a} := \lambda a + (1-\lambda)a'$ and $\tilde{b} := \lambda b + (1-\lambda)b'$. To prove that $(\tilde{a},\tilde{b})^T \in K$ we need to show that $\tilde{a} \geq 1$ and $\tilde{b} \geq \frac{1}{\tilde{a}}$.

Observe that $\tilde{a} = \lambda a + (1-\lambda)a' \geq \min\{\lambda a + (1-\lambda)a, \lambda a' + (1-\lambda)a'\} = 1$ and

$$\frac{1}{\tilde{a}} = \frac{1}{\lambda a + (1-\lambda)a'} \leq \frac{1}{\lambda \frac{1}{b} + (1-\lambda)\frac{1}{b'}} \leq \frac{1}{\lambda \frac{1}{b}} + \frac{1}{(1-\lambda)\frac{1}{b'}} = \frac{1}{\lambda} b + \frac{1}{1-\lambda} b' = \tilde{b}.$$

Thus $(\tilde{a},\tilde{b})^T \in K$ and thus K is convex.

Now let $c := (0,1)^T$. Clearly $\inf\{c^T x \;:\; x \in K\} > -\infty$ since for each $x = (a,b) \in K$ we have $c^T x = b \geq \frac{1}{a} > 0$.

Asume that there is a vector $x^* = (a^*, b^*) \in K$ such that $c^T x^* \leq c^T x$ for all $x \in K$. Since $x^* \in K$, we have $b^* \geq \frac{1}{a^*}$. This implies $b^*/2 \geq \frac{1}{a^*}/2 = \frac{1}{2a^*}$ which shows that $(2a^*, b^*/2)^T \in K$ holds.

Since $b^* > 0$, $c^T x^* = b^* > b^*/2 = c^T(2a^*, b^*/2)^T$, which is a contradiction to the assumption that x^* as defined above exists.

Thus there is no such x^*.

Exercise 4
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map and let $K \subseteq \mathbb{R}^n$ be a set.

1. Show that $f(K) := \{f(x) : x \in K\}$ is convex if K is convex. Is the reverse also true?

2. Prove that $\operatorname{conv}(f(K)) = f(\operatorname{conv}(K))$.

Solution

1. Let K be convex. To prove that $f(K)$ is convex, we need to show that for each two vectors $x, y \in f(K)$ and $\lambda \in [0,1]$ the vector $z := \lambda x + (1-\lambda)y$ is contained in $f(K)$ as well.

Since $x, y \in f(K)$, by definition there are $x^*, y^* \in K$ such that $x = f(x^*)$ and $y = f(y^*)$. Since K is convex we have $z^* := \lambda x^* + (1-\lambda)y^* \in K$. Now

$$f(z^*) = f(\lambda x^* + (1-\lambda)y^*) = \lambda f(x^*) + (1-\lambda)f(y^*) = \lambda x + (1-\lambda)y = z.$$

This shows that $z \in f(K)$ and thus $f(K)$ is convex.

The reverse is false. Consider the linear map

$$f : \mathbb{R}^n \to \mathbb{R}^n \ : \ x \mapsto 0.$$

The set $K := \{(0,\dots,0)^T, (1,\dots,1)^T\} \subset \mathbb{R}^n$ is *not* convex, but $f(K) = \{(0,\dots,0)^T\}$ is.

2. To prove that two sets are equal, it is sufficient to show that each set includes the other.

   We first prove that $\text{conv}(f(K)) \subseteq f(\text{conv}(K))$ holds. Let $x \in \text{conv}(f(K))$. We need to show that $x \in f(\text{conv}(K))$.

   Since $x \in \text{conv}(f(K))$, there are points $x_1,\dots,x_t \in f(K)$ and $\lambda_1,\dots,\lambda_t \geq 0$ such that $\sum_{i=1}^t \lambda_i = 1$ and

$$x = \sum_{i=1}^t \lambda_i x_i.$$

   For each $i = 1,\dots,t$ since $x_i \in f(K)$, by definition there is an $x_i^* \in K$ such that $x_i = f(x_i^*)$. Thus

$$x = \sum_{i=1}^t \lambda_i x_i = \sum_{i=1}^t \lambda_i f(x_i^*) = f\left(\sum_{i=1}^t \lambda_i x_i^*\right) \in f(\text{conv}(K)).$$

   The proof that $\text{conv}(f(K)) \supseteq f(\text{conv}(K))$ is quite similar. Let $x \in f(\text{conv}(K))$. Thus there are points $x_1^*,\dots,x_t^* \in K$ and $\lambda_1,\dots,\lambda_t \geq 0$ such that $\sum_{i=1}^t \lambda_i = 1$ and

$$x = f\left(\sum_{i=1}^t \lambda_i x_i^*\right).$$

   We need to show that $x \in \text{conv}(f(K))$. Let $x_i := f(x_i^*)$ for each $i = 1,\dots,t \in K$. Thus $x_i \in f(K)$. Hence we have

$$x = f\left(\sum_{i=1}^t \lambda_i x_i^*\right) = \sum_{i=1}^t \lambda_i f(x_i^*) = \sum_{i=1}^t \lambda_i x_i \in \text{conv}(f(K)).$$

   Putting both results together we get $\text{conv}(f(K)) = f(\text{conv}(K))$.

Exercise 5
Let $X \subseteq \mathbb{R}^n$ be a set of points. Prove the following statements:

1. The set

$$\text{cone}(X) := \left\{\sum_{i=1}^t \lambda_i x_i \ : \ t \in \mathbb{N}, x_i \in X, \lambda_i \geq 0 \forall i = 1,\dots,t\right\}$$

   is a cone.

2. Each cone containing X also contains $\text{cone}(X)$.

Solution

1. As defined in the lecture, a cone is a set K that is convex and for each $v \in K$ and $\lambda \geq 0$, the vector λv is contained in K as well. Thus to show that $\text{cone}(X)$ is a cone, we have to prove the following:

   a) $\text{cone}(X)$ is convex

   b) For each $v \in \text{cone}(X)$ and $\lambda \in \mathbb{R}_{\geq 0}$ we have $\lambda v \in \text{cone}(X)$.

   To show (a), let $u, v \in \text{cone}(X)$ be two points of the set. We have to show that for each $\lambda \in [0,1]$, the point $\lambda u + (1-\lambda)v$ is contained in $\text{cone}(X)$.

   By definition we have
$$u = \sum_{i=1}^{t} \mu_i x_i \text{ and } v = \sum_{i=1}^{t} \rho_i x_i$$
   for some $x_1, \ldots, x_t \in X$ and $\mu_1, \ldots, \mu_t, \rho_1, \ldots, \rho_t \geq 0$.

   Let $w := \lambda u + (1-\lambda)v$ for some $\lambda \in [0,1]$. We can write
$$\begin{aligned} w = \lambda u + (1-\lambda)v &= \lambda \sum_{i=1}^{t} \mu_i x_i + (1-\lambda) \sum_{i=1}^{t} \rho_i x_i \\ &= \sum_{i=1}^{t} \lambda \mu_i x_i + \sum_{i=1}^{t} (1-\lambda)\rho_i x_i \\ &= \sum_{i=1}^{t} \underbrace{(\lambda \mu_i + (1-\lambda)\rho_i)}_{\geq 0} x_i \end{aligned}$$

   We have written w as linear combination of points of X with nonnegative coefficients which shows that $w \in \text{cone}(X)$ holds. Thus $\text{cone}(X)$ is convex.

   To show (b), let $v \in \text{cone}(X)$ and $\lambda \in \mathbb{R}_{\geq 0}$. By definition we have
$$v = \sum_{i=1}^{t} \mu_i x_i$$
   for some $x_1, \ldots, x_t \in X$ and $\mu_1, \ldots, \mu_t \geq 0$.

   Thus
$$\lambda v = \lambda \sum_{i=1}^{t} \mu_i x_i = \sum_{i=1}^{t} \underbrace{\lambda \mu_i}_{\geq 0} x_i$$

   This shows that $\lambda v \in \text{cone}(X)$, and thus (b) holds.

   We have shown that $\text{cone}(X)$ is a cone.

2. Let C be a cone containing X. We want to show that $\text{cone}(X) \subseteq C$. To do this, we show that for each $t \in \mathbb{N}$, $x_1, \ldots, x_t \in X$ and $\mu_1, \ldots, \mu_t \geq 0$ the vector
$$v = \sum_{i=1}^{t} \mu_i x_i$$

is contained in C.

We do this by induction on t.

For the induction base let $t = 1$. Thus $v = \mu_1 x_1$, where $x_1 \in X$ and $\mu_1 \geq 0$. Since $X \subseteq C$, this implies $x_1 \in C$. As C is a cone, we have

$$C \ni \mu_1 x_1 = v.$$

Our induction hypothesis is that for every $x_1, \ldots, x_t \in X$ and $\mu_1, \ldots, \mu_t \geq 0$ the point

$$\sum_{i=1}^{t} \mu_i x_i$$

is contained in C.

We show that every $x_1, \ldots, x_{t+1} \in X$ and $\mu_1, \ldots, \mu_{t+1} \geq 0$, the point

$$v = \sum_{i=1}^{t+1} \mu_i x_i$$

the point v is contained in C which will complete the proof by induction.

Let v be defined as above. As seen in the base case, $w = \mu_{t+1} x_{t+1} \in C$. We also have $w' := \sum_{i=1}^{t} \mu_i x_i \in C$ by induction hypothesis. Since C is a cone, this implies that $2w \in C$ and $2w' \in C$ holds aswell. Since a cone is convex, we have

$$C \ni \frac{1}{2} 2w + (1 - \frac{1}{2}) 2w' = w + w' = \sum_{i=1}^{t+1} \mu_i x_i = v$$

This ends the proof.