Prof. Friedrich Eisenbrand
Martin Niemeier

Due Date: March 11, 2010
Discussions: February 25, March 04

Discrete Optimization

Spring 2010

Solutions 1

You can hand in written solutions for up to two of the exercises marked with (*) or (Δ) to obtain bonus points. The duedate for this is March 11, 2010, before the exercise session starts. Math students are restricted to exercises marked with (*). Non-math students can choose between (*) and (Δ) exercises.

Exercise 1

A company produces and sells two different products. Our goal is to determine the number of units of each product they should produce during one month, assuming that there is an unlimited demand for the products, but there are some constraints on production capacity and budget.

There are 20000 hours of machine time in the month. Producing one unit takes 3 hours of machine time for the first product and 4 hours for the second product. Material and other costs for producing one unit of the first product amount to 3CHF, while producing one unit of the second product costs 2CHF. The products are sold for 6CHF and 5CHF per unit, respectively. The available budget for production is 4000CHF initially. 25% of the income from selling the first product can be used immediately as additional budget for production, and so can 28% of the income from selling the second product.

1. Formulate a linear program to maximize the profit subject to the described constraints.
2. Solve the linear program graphically by drawing its set of feasible solutions and determining an optimal solution from the drawing.
3. Suppose the company could modernize their production line to get an additional 2000 machine hours for the cost of 400CHF. Would this investment pay off?

Solution

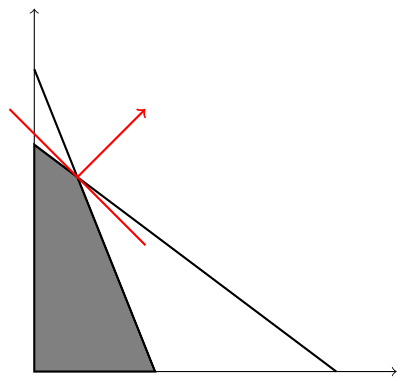
1. Let x be the number of units of the first product and let y be the number of units of the second product.

$$\begin{aligned} \max \quad & (6-3)x + (5-2)y \\ \text{subject to} \quad & 3x + 4y \le 20000 \\ & 3x + 2y \le 4000 + 0.25 \cdot 6x + 0.28 \cdot 5y \\ & x, y \ge 0 \end{aligned}$$

This can be simplified to:

$$\begin{array}{rl} \max & 3x+3y \\ \text{subject to} & 3x+4y \le 20000 \\ & 1.5x+0.6y \le 4000 \\ & x,y \ge 0 \end{array}$$

2. In the following picture, you can see the feasible region and the objective function direction in red.



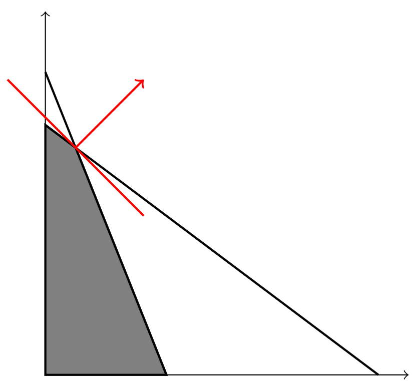
The optimal solution is at the intersection of two of the constraints set to equality, so it is the solution to the system of linear equations:

$$\begin{array}{rcl} 3x+4y & = & 20000 \\ 1.5x+0.6y & = & 4000 \end{array}$$

The unique solution is $1000 \cdot (20/21, 30/7)^T$. Its objective function value is $110000/7 \approx 15714$.

3. Admittedly, there is some ambiguity in the problem statement: It is clear that the first constraint (about the number of machine hours) must be modified, and we will have to subtract the 400CHF from the profit at the end. But should the second constraint (about the available budget) also be changed? The problem statement is not really clear about this. The following assumes that we do not change the second constraint, but the solution is of course similar in the other case as well.

The feasible region looks similar – only one constraint has moved, but not by much:



The optimal solution is now the solution of the linear equations:

$$\begin{aligned} 3x + 4y &= 22000 \\ 1.5x + 0.6y &= 4000 \end{aligned}$$

The unique solution is $1000 \cdot (2/3, 5)^T$. Its objective function value is 17000, so even after subtracting the investment of 400CHF the remaining profit is 16600CHF and thus an improvement over the previous optimum.

Exercise 2
Consider the problem

$$\begin{aligned} \min \quad & 2x + 3|y - 10| \\ \text{subject to} \quad & |x + 2| + y \leq 5, \end{aligned} \tag{1}$$

and reformulate it as a linear programming problem.

Solution
To get a linear program, we need to get rid of the absolute values. We will ensure that our linear program has the following properties.

- Every solution of (1) yields a solution of the linear program with the same objective value.
- Each solution of the linear program yields a solution for (1) with at most that objective value.

If our linear program has these two properties, this implies that optimal solutions for the linear program give optimal solutions for (1) and we are done.

The basic idea is to replace the absolute values with a new variable and two constraints as follows: We replace $|y - 10|$ with a variable z_1 and the constraints $z_1 \geq y - 10$ and $z_1 \geq -y + 10$. I.e. we get

$$\begin{aligned} \min \quad & 2x + 3z_1 \\ \text{subject to} \quad & |x + 2| + y \leq 5 \\ & z_1 \geq y - 10 \\ & z_1 \geq -y + 10. \end{aligned}$$

Note that $z_1 \geq |y - 10|$ in any feasible solution. One can easily check that the two properties from above are fulfilled. Observe that it is crucial at this point that we consider a *minimization* problem. This replacement would not work with a *maximization* problem.

Similarly we replace the term $|x+2|$ with a variable z_2 and two constraints $z_2 \geq x+2$ and $z_2 \geq -x-2$. We get

$$\begin{aligned} \min \quad & 2x + 3z_1 \\ \text{subject to} \quad & z_2 + y \leq 5 \\ & z_1 \geq y - 10 \\ & z_1 \geq -y + 10 \\ & z_2 \geq x + 2 \\ & z_2 \geq -x - 2. \end{aligned}$$

This ensures $z_2 \geq |x+2|$. Again one can easily check that both properties are fulfilled.

Exercise 3
Prove the following statement or give a counterexample: The set of optimal solutions of a linear program is always finite.

Solution
This statement is clearly false. The linear program

$$\max\{x + y \mid x + y \leq 0\}$$

has infinitely many optimal solutions. More generally, if a linear program is bounded, the optimal solutions form a face of the polyhedron of feasible solutions. If this face is at least 1-dimensional, there are infinitely many optimal solutions.

Exercise 4
Let (2) be a linear program in inequality standard form, i.e.

$$\max\{c^T x \mid Ax \leq b, x \in \mathbb{R}^n\} \tag{2}$$

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

Prove that there is an equivalent linear program (3) of the form

$$\min\{\tilde{c}^T x \mid \tilde{A}x = \tilde{b}, x \geq 0, x \in \mathbb{R}^{\tilde{n}}\} \tag{3}$$

where $\tilde{A} \in \mathbb{R}^{\tilde{m}\times\tilde{n}}$, $\tilde{b} \in \mathbb{R}^{\tilde{m}}$, and $\tilde{c} \in \mathbb{R}^{\tilde{n}}$ are such that every optimal point of (2) corresponds to an optimal point of (3) and vice versa.

Linear programs of the form in (3) are said to be in *equality standard form.*

Solution
The transformation requires three steps:

1. Replace every variable x_j with two non-negative variables x_j^+ and x_j^-, and replace every occurrence of x_j with $(x_j^+ - x_j^-)$.

2. Replace every constraint of the form $a_{i1}x_1 + \ldots + a_{in}x_n \leq b_i$ with a constraint $a_{i1}x_1 + \ldots + a_{in} + s_i = b_i$, where s_i is a new, non-negative *slack* variable.

3. Multiply the objective function with -1 to obtain a minimization problem.

Combining these two steps, we can write the transformed linear program as

$$\begin{aligned} \min \quad & -c^T x^+ + c^T x^- \\ \text{subject to} \quad & Ax^+ - Ax^- + s = b \\ & x^+ \geq 0 \\ & x^- \geq 0 \\ & s \geq 0 \end{aligned}$$

This is the desired form if we set $\tilde{c} = \begin{pmatrix} -c & c & 0 \end{pmatrix}$ and $\tilde{A} = \begin{pmatrix} A & -A & I \end{pmatrix}$, where I is the $m \times m$ identity matrix.

Given a feasible solution x of the original linear program, we can find a feasible solution $\tilde{x} = \begin{pmatrix} x^+ & x^- & s \end{pmatrix}^T$ of the reformulated program by setting x^+ to the positive part of x, x^- to the negative part of x, and $s = b - Ax$. It is easy to check that $\tilde{x}$ is feasible and that $\tilde{c}^T \tilde{x} = -c^T x$.

Conversely, given a feasible solution $\tilde{x} = \begin{pmatrix} x^+ & x^- & s \end{pmatrix}^T$ of the reformulated program, it is easy to check that $x = x^+ - x^-$ is a feasible solution of the original linear program with negated objective function value.

This implies in particular that optimal solutions correspond to each other.

Exercise 5

Recall the image decomposition problem for OLEDs from the lecture and its formulation as a linear program :

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} u_i^{(1)} + \sum_{i=1}^{n-1} u_i^{(2)} && && (4) \\ \text{s.t.} \quad & f_{ij}^{(1)} + f_{i-1,j}^{(2)} + f_{ij}^{(2)} = r_{ij} && \text{for all } i, j \\ & f_{ij}^{(\alpha)} \leq u_i^{(\alpha)} && \text{for all } i, j, \alpha \\ & f_{ij}^{(\alpha)} \geq 0 && \text{for all } i, j, \alpha \end{aligned}$$

Apply this technique to find an optimal decomposition of the EPFL logo[1] using Zimpl and LP solver libraries:

1. Familiarize yourself with the Zimpl modelling language[2].

2. Model the linear program (4) to decompose the EPFL logo with Zimpl. You can find an incomplete model containing the encoding of the grayscale values of the logo here[3].

[1] http://www.epfl.ch/images/EPFL-logo.jpg
[2] http://www.zib.de/koch/zimpl/download/zimpl.pdf
[3] http://disopt.epfl.ch/webdav/site/disopt/users/190205/public/logo_dec.zmpl

3. Solve the linear program using an LP-solver like QSopt[4], lp_solve[5] or SoPlex[6].

Solution

The linear program (4) can be modelled with zimpl as follows (The zimpl file can be downloaded here[7]):

```
# The image to decompose
set rows := { 1 to 34};
set columns := { 1 to 120};
param r[rows * columns ] := <1,1> 87, <1,2> 87, ... , <34,120> 87;

# The variables of the linear program
var u1[rows] >= -infinity <= infinity;
var u2[rows] >= -infinity <= infinity;
var f1[rows * columns] >= 0 <= infinity;
var f2[rows * columns] >= 0 <= infinity;

# The objective function
minimize cost: (sum <i> in rows : u1[i]) +  (sum <i> in rows : u2[i]);

# Decomposition constraints
subto dec1: forall<i,j> in rows * columns with i>1 do
    f1[i,j]+f2[i-1,j]+f2[i,j]==r[i,j];
# Special constraint for the first row
subto dec2: forall<j> in columns  do
    f1[1,j]+f2[1,j]==r[1,j];

# bounds for the u-variables
subto bound1: forall<i,j> in rows * columns do
    f1[i,j]<=u1[i];
subto bound2: forall<i,j> in rows * columns do
    f2[i,j]<=u2[i];
```

We can convert the zimpl file to an mps file by typing:

```
zimpl -t mps logo\_dec\_sol.zmpl
```

This generates the file *logo_dec_sol.mps*, which can be read by lp solver libraries such as *lp_solve*:

```
lp_solve -mps logo\_dec\_sol.mps
```

[4]http://www2.isye.gatech.edu/ wcook/qsopt/
[5]http://lpsolve.sourceforge.net/5.5/
[6]http://soplex.zib.de/
[7]http://disopt.epfl.ch/webdav/site/disopt/users/190205/public/assignments/logo_dec_sol.zmpl

This gives an optimal solution to the linear program. The value of the objective function is 4649.5. Thus we have found a decomposition of the image that needs time 4649.5 for display. The traditional approach takes time 8670.

Exercise 6 (*)

1. Let $\{C_i\}_{i\in I}$ be a family of convex subsets of $\mathbb{R}^n$. Prove that the intersection $\bigcap_{i\in I} C_i$ is convex.
2. Let $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. Prove that the *closed halfspace* $H := \{x \in \mathbb{R}^n \mid a^T x \leq b\}$ is closed and convex.
3. Prove that the set of feasible points of a linear program is convex and closed.
4. Find a convex and closed set $C \subset \mathbb{R}^2$ and a vector $c \in \mathbb{R}^2$ such that $\sup\{c^T x \mid x \in C\}$ is finite but $\max\{c^T x \mid x \in C\}$ does not exist.

Solution

1. Let $u, v \in \bigcap_{i\in I} C_i$. We need to show that for all $\lambda \in [0,1]$, the point $\lambda u + (1-\lambda)v \in \bigcap_{i\in I} C_i$. So let $\lambda \in [0,1]$. Note that, since each C_i is convex, we have $\lambda u + (1-\lambda)v \in C_i$ for all $i \in I$. Hence $\lambda u + (1-\lambda)v \in \bigcap_{i\in I} C_i$, completing the proof.
2. *H is closed.* This can be shown directly using the definition of closed sets. Alternatively, observe that the complement of H is the preimage of the open interval (b,∞) under the linear (and thus continuous) map $f(x) := a^T x$. Therefore the complement of H is open, and H is closed.

   H is convex. Let $u, v \in H$ and $0 \leq \lambda \leq 1$. We need to show that the point $p = \lambda u + (1-\lambda)v \in H$. To do this, simply calculate:

   $$a^T p = a^T(\lambda u + (1-\lambda v)) = \lambda(a^T u) + (1-\lambda)(a^T v) \leq \lambda b + (1-\lambda)b = b$$

3. For any linear program, we can write the set P of feasible points as $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$, where $A \in \mathbb{R}^{m\times n}$ is a matrix and $b \in \mathbb{R}^m$ is a vector. Let $a_1, \ldots, a_m$ be the rows of A and $b_1, \ldots, b_m$ the components of b. Then we can write

   $$P = \{x \in \mathbb{R}^n \mid a_j x \leq b_j \text{ for all } j = 1\ldots m\} = \bigcap_{j=1}^{m} \{x \in \mathbb{R}^n \mid a_j x \leq b_j\} = \bigcap_{j=1}^{m} H_j$$

   where $H_j := \{x \in \mathbb{R}^n \mid a_j x \leq b_j\}$ is the closed halfspace determined by the inequality $a_j x \leq b_j$.

   Each H_j is closed and convex by the previous part, and we know that a finite intersection of such sets is still closed and convex.

4. Consider, for example, the set of points above the graph of the exponential function:

$$C := \{\begin{pmatrix} x \\ y \end{pmatrix} \mid y \geq e^x\}$$

C is closed. The complement of C can be written as $f^{-1}((-\infty, 0))$, where f is the continuous function $f(\begin{pmatrix} x \\ y \end{pmatrix}) := y - e^x$. Thus the complement of C is open.

C is convex. It is known from analysis that the exponential function is convex, that is for all $x_1, x_2 \in \mathbb{R}$, $\lambda \in [0,1]$ we have $\lambda e^{x_1} + (1-\lambda)e^{x_2} \geq e^{\lambda x_1 + (1-\lambda)x_2}$.

Now set $c := \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Clearly, $\sup\{c^T p \mid p \in C\} = 0$, because C lies above the x-axis and $\lim_{x\to-\infty} e^x = 0$. On the other hand, no point of C lies on the x-axis because the exponential function is strictly positive, so no maximum exists.

Exercise 7 (Δ)
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 $K := \{(0,\ldots,0)^T, (1,\ldots,1)^T\} \subset \mathbb{R}^n$ is *not* convex, but $f(K) = \{(0,\ldots,0)^T\}$ is.

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

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

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 8 (*)
Consider the vectors

$$x_1 = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, x_2 = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, x_3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, x_4 = \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}, x_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The vector

$$v = x_1 + 3x_2 + 2x_3 + x_4 + 3x_5 = \begin{pmatrix} 15 \\ 14 \\ 25 \end{pmatrix}$$

is a conic combination of the x_i.
Write v as a conic combination using only three vectors of the x_i.

Hint: Recall the proof of Carathéodory's theorem

Solution
Let $X = \{x_1, \dots, x_n\}$. Observe that $v \in \text{cone}(X)$. Since $x_1, \dots, x_5 \in \mathbb{R}^3$, Carathéodory's theorem states that we can write v as a conic combination using at most three vectors of X.

How to compute this conic combination? Recall the proof of Carathéodory's theorem. The number of vectors in the conic combination $v = \sum_{i=1}^{5} \lambda_i x_i$ can be reduced by one with the following method: Compute a nontrivial linear combination of the all zero vector, i.e. compute $\mu_1, \dots, \mu_5 \in \mathbb{R}$, not all of them zero such that $\sum_{i=1}^{5} \mu_i x_i = 0$ holds.

Thus $v = \sum_{i=1}^{5} (\lambda_i - \epsilon\mu_i) x_i$ for each $\epsilon > 0$. As described in the proof, one can find an ϵ^* such that $\lambda_i - \epsilon^* \mu_i \geq 0$ for each $i = 1, \dots, 5$ and $\lambda_i - \epsilon^* \mu_i = 0$ for at least one i. Thus we get a new conic combination of v using one vector less than before.

We now applying the idea to the exercise. We first compute a nontrivial linear combination of the all zero vector by solving the following system of linear equations:

$$\begin{aligned} 3\mu_1 + \mu_2 + 2\mu_3 + 2\mu_4 + \mu_5 &= 0 \\ \mu_1 + 2\mu_2 + 4\mu_4 + \mu_5 &= 0 \\ 2\mu_1 + 5\mu_2 + \mu_3 + 3\mu_4 + \mu_5 &= 0. \end{aligned}$$

Using standard methods, e.g. gaussian elimination, one can compute the solution set as:

$$S = \{(-4a - b, 0, 5a + b, a, b) \; : \; a, b \in \mathbb{R}\}$$

We take the solution $(-5, 0, 6, 1, 1) \in S$ which gives a nontrivial linear combination, i.e. $0 = -5x_1 + 6x_3 + x_4 + x_5$.

What is the maximal ϵ such that

$$v = (1 + 5\epsilon)x_1 + 3x_2 + (2 - 6\epsilon)x_3 + (1 - \epsilon)x_4 + (3 - \epsilon)x_5$$

is a conic combination? Each coefficient has to be nonnegative, thus observe that

$$\epsilon^* = \frac{1}{3}$$

is the maximum. We get the new conic combination

$$v = \frac{8}{3}x_1 + 3x_2 + 0x_3 + \frac{2}{3}x_4 + \frac{8}{3}x_5.$$

Observe that since the coefficient of x_3 is zero, we can remove it from the conic combination.

We need to remove one more vector to get a conic combination using only three vectors. Again we compute a nontrivial linear combination of the all zero vector using the remaining vectors (x_1, x_2, x_4, x_5):

$$\begin{aligned} 3\mu_1 + \mu_2 + 2\mu_4 + \mu_5 &= 0 \\ \mu_1 + 2\mu_2 + 4\mu_4 + \mu_5 &= 0 \\ 2\mu_1 + 5\mu_2 + 3\mu_4 + \mu_5 &= 0. \end{aligned}$$

We compute the solution set which is

$$S' = \{(-a, 0, -a, 5a) \; : \; a \in \mathbb{R}\}$$

We take the solution $(-1, 0, -1, 5) \in S'$ which gives a nontrivial linear combination, i.e. $0 = -x_1 - x_3 + 5x_5$.

What is the maximal ϵ such that

$$v = (\frac{8}{3} + \epsilon)x_1 + 3x_2 + (\frac{2}{3} + \epsilon)x_4 + (\frac{8}{3} - 5\epsilon)x_5$$

is a conic combination? It is given by $\epsilon^* = \frac{8}{15}$.

The new conic combination is

$$\begin{aligned} v &= (\frac{8}{3} + \frac{8}{15})x_1 + 3x_2 + (\frac{2}{3} + \frac{8}{15})x_4 + (\frac{8}{3} - 5\frac{8}{15})x_5 \\ &= \frac{16}{5}x_1 + 3x_2 + \frac{6}{5}x_4 + 0x_5. \end{aligned}$$

Since the coefficient of x_5 is zero, we can remove it and obtain the desired convex combination of v using only three vectors.