Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Mohsen Ghaffari, Angelika Steger, Emo Welzl, Peter Widmayer

Algorithms, Probability, and Computing **Solutions KW47** **HS16**

Solution for the in-class exercise

As in the lecture notes we use slack variables to get (3) into form (1):

$$\begin{array}{lrcl} \text{maximize} & c^T(x_1 - x_2) & & \\ \text{subject to} & A(x_1 - x_2) & \leq & b \\ & x_1, x_2 & \geq & 0 \end{array} \tag{1}$$

This is the same as

$$\begin{array}{lrcl} \text{maximize} & \begin{pmatrix} c \\ -c \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & & \\ \text{subject to} & (A, -A) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \leq & b \\ & \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \geq & 0. \end{array} \tag{2}$$

Using the definition of the dual we get that the dual of (3) is

$$\begin{array}{lrcl} \text{minimize} & b^T y & & \\ \text{subject to} & (A, -A)^T y & \geq & \begin{pmatrix} c \\ -c \end{pmatrix} \\ & y & \geq & 0. \end{array} \tag{3}$$

The claim follows since

$$(A, -A)^T y \geq \begin{pmatrix} c \\ -c \end{pmatrix}$$

is equivalent to $A^T y = c$.

Solution 1: Solving Linear Programs via Binary Search

Suppose we are given a linear program L in the form

$$\text{maximize } c^T x \text{ subject to } Ax \leq b.$$

Suppose that L has an optimal solution $\mathbf{x}^*$ with objective value $\text{OPT} := \mathbf{c}^T\mathbf{x}^*$. By theorem 6.2 we can assume $\mathbf{x}^* \in [-K, K]^n$ with $K \leq 2^{O(\langle L \rangle)}$. The theorem also tells us $\langle \mathbf{x}_j^* \rangle = O(\langle L \rangle)$, from which we can deduce that also $\langle \text{OPT} \rangle \in O(\langle L \rangle)$ (we skip the calculation here). Thus $\langle \text{OPT} \rangle \leq C \cdot \langle L \rangle$ for some absolute constant $C > 0$.

Our algorithm must actually know K and C explicitly in order to proceed, but we consider the exact values an implementation detail.

It is an easy task to find a vertex $\mathbf{x}_{\max}$ (resp., $\mathbf{x}_{\min}$) of the cube $[-K, +K]^n$ which maximizes $\mathbf{c}^T\mathbf{x}$ (resp., minimizes $\mathbf{c}^T\mathbf{x}$). Indeed, the sign of any coordinate of $\mathbf{c}$ corresponds to the sign of the corresponding coordinate of $\mathbf{x}_{\max}$, and once we have found $\mathbf{x}_{\max}$ we can choose $\mathbf{x}_{\min} = -\mathbf{x}_{\max}$. Since, as we said earlier, $\mathbf{x}^*$ is contained in $[-K, K]^n$ we get that $\alpha := \mathbf{c}^T\mathbf{x}_{\max}$ and $\beta := \mathbf{c}^T\mathbf{x}_{\min}$ are upper and lower bounds, respectively, for OPT.

So far we have assumed that L has an optimal solution; but first, how can we check if this is true? Otherwise either (1) L is infeasible, or (2) the objective function is unbounded. Case (1) we can easily check with one call to our feasibility oracle. Case (2) we can rule out by adding a constraint $c^T x \geq \alpha + 1$ and checking if the resulting linear program is still feasible.

Now we can perform binary search for OPT. That is, we let $\gamma := \frac{1}{2}(\alpha + \beta)$ and we add the constraint $\mathbf{c}^T\mathbf{x} \geq \gamma$ to L. We check whether the new program is still feasible. If it is, then we update the lower bound $\beta := \gamma$. If it is not, then we remove the new constraint again and we update the upper bound $\alpha := \gamma$. In any case, the size of the interval $[\beta, \alpha]$ that contains OPT halves in every step of the search. We stop the binary search as soon as the set

$$[\beta, \alpha] \cap \{q \in \mathbb{Q} : \langle q \rangle \leq C \cdot \langle L \rangle\} \tag{$*$}$$

contains only one element, which must then be equal to OPT. In a last step we find a feasible solution to L with the additional constraint $c^T x = \text{OPT}$.

How do we check if $(*)$ contains only one element? **Claim:** For any two rational numbers $a_1, a_2 \in \mathbb{Q}$, if $a_1 \neq a_2$ and $\langle a_1 \rangle, \langle a_2 \rangle \leq \zeta$ for some $\zeta > 0$, then $|a_1 - a_2| \geq \frac{1}{2^{2\zeta}}$. – Indeed, if we write $a_i = \frac{p_i}{q_i}$ with $p_i \in \mathbf{Z}$, $q_i \in \mathbf{N}$, then

$$|a_1 - a_2| \geq \frac{1}{q_1 q_2} \geq \frac{1}{2^{\langle q_1 \rangle} 2^{\langle q_2 \rangle}} \geq \frac{1}{2^{\langle a_1 \rangle + \langle a_2 \rangle}} \geq \frac{1}{2^{2\zeta}}.$$

From the claim it follows that $(*)$ is a singleton set as soon as the size of the interval $[\beta, \alpha]$ is less than $(1/2)^{2C\langle L \rangle}$; so this will be our termination criterion.

The number of steps until we reach our termination criterion is at most

$$\log_2 \left(\frac{\alpha - \beta}{(1/2)^{2C \cdot \langle L \rangle}} \right) \leq \log_2(2^{O(\langle L \rangle)}) \leq O(\langle L \rangle)$$

where we have used that $|\alpha|, |\beta| \leq 2^{O(\langle L \rangle)}$ (again we skip the calculation).

For our method to make sense it remains to check that the auxiliary linear programs that we used all have size polynomial in $\langle L \rangle$. We skip the details.

Solution 2: Equivalence of the Three Farkas Lemmas

(a) This is almost the exact same argument as in the previous exercise. Suppose there is $\mathbf{y}$ with $A^\mathsf{T}\mathbf{y} \geq \mathbf{0}$ and $\mathbf{b}^\mathsf{T}\mathbf{y} < 0$. Furthermore, towards a contradiction, suppose there is an $\mathbf{x} \geq \mathbf{0}$ with $A\mathbf{x} = \mathbf{b}$ (or, equivalently, $\mathbf{x}^\mathsf{T}A^\mathsf{T} = \mathbf{b}^\mathsf{T}$). We arrive at a contradiction (and hence conclude that $A\mathbf{x} = \mathbf{b}$ has no non-negative solution) by observing that

$$0 \leq \mathbf{x}^\mathsf{T}A^\mathsf{T}\mathbf{y} = \mathbf{b}^\mathsf{T}\mathbf{y} < 0,$$

where the first inequality is justified because both $\mathbf{x}^\mathsf{T}$ and $A^\mathsf{T}\mathbf{y}$ are non-negative.

Now suppose instead that there is $\mathbf{y} \geq \mathbf{0}$ with $A^\mathsf{T}\mathbf{y} \geq \mathbf{0}$ and $\mathbf{b}^\mathsf{T}\mathbf{y} < 0$. Furthermore, towards a contradiction, suppose there is an $\mathbf{x} \geq \mathbf{0}$ with $A\mathbf{x} \leq \mathbf{b}$ (or, equivalently, $\mathbf{x}^\mathsf{T}A^\mathsf{T} \leq \mathbf{b}^\mathsf{T}$). We arrive at a contradiction (and hence conclude that $A\mathbf{x} \leq \mathbf{b}$ has no non-negative solution) by observing that

$$0 \leq \mathbf{x}^\mathsf{T}A^\mathsf{T}\mathbf{y} \leq \mathbf{b}^\mathsf{T}\mathbf{y} < 0,$$

where the second inequality is justified because of our assumption $A\mathbf{x} \leq \mathbf{b}$ and because $\mathbf{y}$ is non-negative.

(b) We only prove the implication I $\Rightarrow$ II. The other implications can be proved in a very similar fashion.

We note that $A\mathbf{x} = \mathbf{b}$ has no non-negative solution $\mathbf{x}$ if and only if the system $\mathcal{A}\mathbf{x} \leq \mathcal{B}$ is unsolvable, where

$$\mathcal{A} := \begin{pmatrix} A \\ -A \\ -\mathbb{1} \end{pmatrix}, \qquad \mathcal{B} := \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \\ \mathbf{0} \end{pmatrix},$$

and where $\mathbb{1}$ is the identity matrix of appropriate dimension. Indeed, the system $\mathcal{A}\mathbf{x} \leq \mathcal{B}$ just encodes the constraints $A\mathbf{x} \leq \mathbf{b}$, $A\mathbf{x} \geq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. We now use Farkas lemma I and see that $\mathcal{A}\mathbf{x} \leq \mathcal{B}$ is unsolvable if and only if there is a vector $\mathcal{Y} \geq \mathbf{0}$ with $\mathcal{A}^\mathsf{T}\mathcal{Y} = \mathbf{0}$ and $\mathcal{B}^\mathsf{T}\mathcal{Y} < 0$. We write the vector $\mathcal{Y}$ as

$$\mathcal{Y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \end{pmatrix}$$

so that $\mathcal{A}^\mathsf{T}\mathcal{Y} = A^\mathsf{T}\mathbf{y}_1 - A^\mathsf{T}\mathbf{y}_2 - \mathbf{y}_3$ and $\mathcal{B}^\mathsf{T}\mathcal{Y} = \mathbf{b}^\mathsf{T}\mathbf{y}_1 - \mathbf{b}^\mathsf{T}\mathbf{y}_2$. Finally, we note that there exists $\mathcal{Y} \geq \mathbf{0}$ with $\mathcal{A}^\mathsf{T}\mathcal{Y} = \mathbf{0}$ and $\mathcal{B}^\mathsf{T}\mathcal{Y} < 0$ if and only if there is $\mathbf{y}$ with $A^\mathsf{T}\mathbf{y} \geq 0$ and $\mathbf{b}^\mathsf{T}\mathbf{y} < 0$, which concludes the proof. Indeed, for the "only if" we define $\mathbf{y} := \mathbf{y}_1 - \mathbf{y}_2$ and see that

$$A^\mathsf{T}\mathbf{y} = A^\mathsf{T}\mathbf{y}_1 - A^\mathsf{T}\mathbf{y}_2 \geq A^\mathsf{T}\mathbf{y}_1 - A^\mathsf{T}\mathbf{y}_2 - \mathbf{y}_3 = \mathcal{A}^\mathsf{T}\mathcal{Y} = \mathbf{0},$$

where the inequality is justified because $\mathcal{Y}$ (and thus also $\mathbf{y}_3$) is non-negative, and we also see that

$$\mathbf{b}^\mathsf{T}\mathbf{y} = \mathbf{b}^\mathsf{T}\mathbf{y}_1 - \mathbf{b}^\mathsf{T}\mathbf{y}_2 = \mathcal{B}^\mathsf{T}\mathcal{Y} < \mathbf{0}.$$

For the "if" we can always choose $\mathbf{y}_1$ and $\mathbf{y}_2$ in such a way that both are non-negative and such that $\mathbf{y}_1 - \mathbf{y}_2 = \mathbf{y}$. Since we know that $A^\mathsf{T}\mathbf{y} = A^\mathsf{T}\mathbf{y}_1 - A^\mathsf{T}\mathbf{y}_2 \geq \mathbf{0}$ we can also choose a non-negative $\mathbf{y}_3$ with $A^\mathsf{T}\mathbf{y} - A^\mathsf{T}\mathbf{y}_2 - \mathbf{y}_3 = \mathbf{0}$.

Solution 3: Deciding Feasibility vs. Finding Feasible Solutions

First, we check with one call to the oracle whether the given system of linear inequalities has a solution. If the answer is NO then we stop and also output NO. If the answer is YES then we proceed as follows.

(a) If there are only equations in the system, then this just means that we have a system $A\mathsf{x} = \mathbf{b}$ of linear equations, for which we can find a solution in polynomial time by Gauss elimination.[1] So, in this case we need no additional calls to the oracle.

(b) If there is at least one inequality, say $\mathsf{ax} \leq b$, then we replace it by $\mathsf{ax} = b$. If the new, more constrained, system still has a solution (which can be checked with one additional call to the oracle), then we can recursively find a solution to the original system by finding a solution to the more constrained system. If the new, more constrained, system turns out to have no solution, then we drop the constraint $\mathsf{ax} \leq b$ completely to obtain a smaller system, which again can be solved recursively. This is a sound strategy because replacing a constraint $\mathsf{ax} \leq b$ with $\mathsf{ax} = b$ can turn a feasible problem into an infeasible one if and only if the hyperplane corresponding to $\mathsf{ax} \leq b$ is not part of the boundary of the feasible region (in other words, it is redundant).

Since we need one initial call to the oracle, and exactly one call per inequality that we get rid of (either by replacing it with an equality or by dropping it completely), the total number of calls to the oracle will be $m+1$, where m is the number of inequalities in the original system.

[1] When the computation has to be done exactly, naive implementations of Gauss elimination can lead to an exponential blow-up of the encoding size of intermediate results. However, there are more clever implementations which do not have this problem and which do run in polynomial time.