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 KW46** **HS16**

Solution 1: Finding a Separating Line

Our linear program has variables a, b, c, ε and looks as follows (the number of constraints is $|R| + |B| + 1$):

$$\begin{aligned} \text{maximize } \varepsilon \text{ subject to } & ax + by \geq c + \varepsilon \quad \text{(for each } (x, y) \in R\text{)}, \\ & ax + by \leq c - \varepsilon \quad \text{(for each } (x, y) \in B\text{)}, \\ & \varepsilon \leq 1. \end{aligned}$$

We will show the following (assuming $R, B \neq \emptyset$):

1. Our LP has an optimal solution.
2. If some optimal solution $(a^*, b^*, c^*, \varepsilon^*)$ satisfies $\varepsilon^* > 0$ and $(a^*, b^*) \neq (0, 0)$, then there is a seperating line, namely the line $a^*x + b^*y = c^*$.
3. Conversely, if there is a separating line, then every optimal solution $(a^*, b^*, c^*, \varepsilon^*)$ satisfies $\varepsilon^* > 0$ and $(a^*, b^*) \neq (0, 0)$.

From this it will be clear how to decide, given an optimal solution, if a separating line exists (namely, check if $\varepsilon^* > 0$ and $(a^*, b^*) \neq (0, 0)$) and, if so, how to compute one (take the line $a^*x + b^*y = c^*$).

Ad 1: Our LP is feasible, because the origin is a feasible point. Furthermore the objective function is bounded, because we artificially bounded it with the constraint $\varepsilon \leq 1$. Hence our LP has an optimal solution.

Ad 2: Clear, because by construction any feasible point (a, b, c, ε) with $\varepsilon > 0$ and $(a, b) \neq (0, 0)$ defines a separating line $ax + by = c$.

Ad 3: Assume there exists a separating line $\ell : a_0x + b_0y = c_0$, where we choose the coefficients in a normalized way such that $a_0^2 + b_0^2 = 1$ and such that $a_0x + b_0y > c_0$ for all $(x, y) \in R$. Let $\varepsilon_0 := \min\{1, \text{dist}(\ell, R), \text{dist}(\ell, B)\}$. Then $(a_0, b_0, c_0, \varepsilon_0)$ is feasible with $\varepsilon_0 > 0$. This shows that the optimal value (which we already know to exist) is positive: $\varepsilon^* \geq \varepsilon_0 > 0$. It remains to show $(a^*, b^*) \neq (0, 0)$. Assume otherwise. Then $c^* + \varepsilon^* \leq 0 \leq c^* - \varepsilon^*$ implies $\varepsilon^* \leq 0$, a contradiction to what we proved a moment ago.

Solution 2: Fitting a Ball into a Convex Polytope

We assume that the given intersection of halfspaces is bounded, so that the question is well-defined and fits the title of the exercise. (This was an oversight in phrasing the question.)

We may assume without loss of generality that every constraint $\mathbf{a}_i^T\mathbf{x} \leq b_i$ (which corresponds to the halfspace H_i) is normalized, which simply means that we have $||\mathbf{a}_i||_2 = 1$. Indeed, this can always be achieved by rescaling all the coefficients in a given constraint.

From linear algebra we recall that $\mathbf{a}_i$ is nothing else than the normal vector of the hyperplane P_i defined by the equation $\mathbf{a}_i^T\mathbf{x} = b_i$ (note that P_i is simply the boundary of H_i). Furthermore, the absolute value of b_i is equal the Euclidean distance between P_i and the origin (note that this is only true because we have normalized constraints). The sign of b_i additionally tells us on which side of the coordinate system P_i lies relative to $\mathbf{a}_i$. If $b_i = 0$ then P_i goes through the origin. If $b_i > 0$ then P_i has been moved away from the origin in the direction of the normal vector $\mathbf{a}_i$. If $b_i < 0$ then P_i has been moved away from the origin against the direction of $\mathbf{a}_i$.

Let us now define for every halfspace H_i another halfspace $H_i^r := \{\mathbf{x}\colon \mathbf{a}_i^T\mathbf{x} \leq b_i - r\}$ for a non-negative parameter which we call r. We also define the corresponding hyperplanes P_i^r. Note that for $r > 0$, H_i^r is smaller than H_i in the sense that it is a strict subset. More precisely, the boundary P_i^r of H_i^r has been moved inwards by a distance of r when compared with the boundary P_i of H_i.

Now suppose $r \geq 0$ and that $H^r := \bigcap_{i=1}^m H_i^r$ is non-empty. Fix any point $\mathbf{c}$ in H^r. Clearly, the ball with center point $\mathbf{c}$ and radius r must be completely contained in $H := \bigcap_{i=1}^m H_i$ because $\mathbf{c}$ has distance at least r from P_i for all indices i. Conversely, for any given ball with center point $\mathbf{c}$ and radius r that is completely contained in H, it must also be the case that $\mathbf{c}$ is contained in H_r for otherwise the given ball would properly intersect one of the hyperplanes P_i. Therefore, the desired largest radius r^* is equal to the largest value of $r \geq 0$ with $H^r \neq \emptyset$, and the desired center point $\mathbf{c}^*$ can be any point in H^{r^*}. All of the above conditions can easily be expressed in the following linear program with real variables $\mathbf{c} \in \mathbf{R}^n$ and $r \in \mathbf{R}$.

$$
\begin{aligned}
\text{Maximize} \quad & r \\
\text{subject to} \quad & r \geq 0 \\
& \mathbf{a}_1^T\mathbf{c} \leq b_1 - r \\
& \mathbf{a}_2^T\mathbf{c} \leq b_2 - r \\
& \vdots \\
& \mathbf{a}_m^T\mathbf{c} \leq b_m - r
\end{aligned}
$$

Solution 3: Linear Programs in Equational Form

Given a linear program in standard form[1],

$$\text{maximize } c^T x \text{ subject to } Ax \le b, \tag{LP 1}$$

where $A \in \mathbf{R}^{m \times n}$, $c \in \mathbf{R}^n$ and $b \in \mathbf{R}^m$, we can convert it into equational form as follows: First we replace the "$\le$" by a "$=$" by the following trick.

$$\text{maximize } c^T x \text{ subject to } Ax + \varepsilon = b,\ \varepsilon \ge 0 \tag{LP 2}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$ is a vector of m new variables. In a second step we get rid of unconstrained (= possibly negative) variables. To this end we replace each x_i by $x_i' - x_i''$, where x_i', x_i'' are two new nonnegative variables:

$$\text{maximize } c^T(x' - x'') \text{ subject to } A(x' - x'') + \varepsilon = b,\ x' \ge 0,\ x'' \ge 0,\ \varepsilon \ge 0. \tag{LP 3}$$

Now we write this LP in such a way that it is undoubtedly in equational form, e. g. like this:

$$\text{maximize } \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix}^T \begin{pmatrix} x' \\ x'' \\ \varepsilon \end{pmatrix} \text{ subject to } \begin{pmatrix} A & -A & \mathbf{I}_m \end{pmatrix} \begin{pmatrix} x' \\ x'' \\ \varepsilon \end{pmatrix} = b, \begin{pmatrix} x' \\ x'' \\ \varepsilon \end{pmatrix} \ge 0. \tag{LP 4}$$

In what sense have we "converted" the LP into equational form? We make the following notes.

- If x is a feasible solution of (LP 1), then a corresponding feasible solution of (LP 4) is given by

$$\begin{aligned} x' &= (\max\{x_1, 0\}, \ldots, \max\{x_n, 0\}), \\ x'' &= -(\min\{x_1, 0\}, \ldots, \min\{x_n, 0\}), \\ \varepsilon &= b - Ax. \end{aligned}$$

  Furthermore this feasible solution to (LP 4) has the same objective value as x.

- Correspondingly, if (x', x'', ε) is a feasible solution of (LP 4), then $x' - x''$ is a feasible solution of (LP 1) with the same objective value.

Complexity: The linear program (LP 4) has $2n + m$ variables and $2n + 2m$ constraints, where the original (LP 1) had n variables and m constraints.

[1] You might want to revise at this point how to convert any linear program into standard form (section 6.1 of the lecture notes).

Solution 4: Maximum Number of Vertices of 3-dimensional Convex Polytopes

The vertex-edge graph of P is a planar graph. (Why? Think of cutting a small hole into some face of P and then stretching the whole thing flat.) For the number of faces, f, we have $f \leq n$ because each of our halfspaces defines at most one (unique) face of the polytope. In order to bound the number of edges, e, we count the vertex-edge incidences in two ways:

$$3v \leq \#\left\{(\xi, \eta) \ : \ \xi \text{ is a vertex, } \eta \text{ is an edge that is incident to } \xi\right\} = 2e.$$

Thus $e \geq \frac{3}{2}v$. Plugging this into Euler's formula we find $2 = v - e + f \leq v - \frac{3}{2}v + f = f - \frac{v}{2}$, which gives $v \leq 2f - 4 \leq 2n - 4$.

Solution 5: Certificates for Infeasibility of Systems of Linear Equations

Let us first do the easy direction. Suppose there is $\mathbf{y}$ with $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$ and $\mathbf{b}^{\mathsf{T}}\mathbf{y} = 1$. Furthermore, towards a contradiction, suppose there is an $\mathbf{x}$ 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 $A\mathbf{x} = \mathbf{b}$ is unsolvable) by observing that

$$0 = \mathbf{x}^{\mathsf{T}}\mathbf{0} = \mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}\mathbf{y} = \mathbf{b}^{\mathsf{T}}\mathbf{y} = 1.$$

For the other direction we recall some notation from linear algebra (we assume throughout that the matrix A has m rows and n columns). The *image* of A is the set $\operatorname{img}(A) := \{A\mathbf{x} \mid \mathbf{x} \in \mathbf{R}^n\}$. The *left nullspace* (or *cokernel*) of A is the set $\ker(A^{\mathsf{T}}) := \{\mathbf{y} \in \mathbf{R}^m \mid A^{\mathsf{T}}\mathbf{y} = \mathbf{0}\}$. We also recall that these two sets are vector spaces and that they are orthogonal complements of each other. In particular, if $\mathbf{i}_1, \ldots, \mathbf{i}_r$ is an orthonormal basis of $\operatorname{img}(A)$ and $\mathbf{k}_1, \ldots, \mathbf{k}_s$ is an orthonormal basis of $\ker(A^{\mathsf{T}})$, then $\mathbf{i}_1, \ldots, \mathbf{i}_r, \mathbf{k}_1, \ldots, \mathbf{k}_s$ is an orthonormal basis of $\mathbf{R}^m$.

Now suppose that the system $A\mathbf{x} = \mathbf{b}$ is unsolvable. We show how to construct $\mathbf{y}$ with $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$ and $\mathbf{b}^{\mathsf{T}}\mathbf{y} = 1$. First we write $\mathbf{b}$ as a linear combination $\mathbf{b} = \alpha_1\mathbf{i}_1 + \cdots + \alpha_r\mathbf{i}_r + \beta_1\mathbf{k}_1 + \cdots + \beta_s\mathbf{k}_s$. We observe that $s \geq 1$ and that for some index i we must have $\beta_i \neq 0$ (for otherwise $\mathbf{b} \in \operatorname{img}(A)$, which cannot be if $A\mathbf{x} = \mathbf{b}$ is unsolvable). W.l.o.g. we assume that $\beta_1 \neq 0$ and we define $\mathbf{y} := \frac{1}{\beta_1}\mathbf{k}_1$. We now see that $A^{\mathsf{T}}\mathbf{y} = \mathbf{0}$ because $\mathbf{y} \in \ker(A^{\mathsf{T}})$. Moreover,

$$\mathbf{b}^{\mathsf{T}}\mathbf{y} = \frac{\alpha_1}{\beta_1}\underbrace{\mathbf{i}_1^{\mathsf{T}}\mathbf{k}_1}_{=0} + \cdots + \frac{\alpha_r}{\beta_1}\underbrace{\mathbf{i}_r^{\mathsf{T}}\mathbf{k}_1}_{=0} + \frac{\beta_1}{\beta_1}\underbrace{\mathbf{k}_1^{\mathsf{T}}\mathbf{k}_1}_{=1} + \frac{\beta_2}{\beta_1}\underbrace{\mathbf{k}_2^{\mathsf{T}}\mathbf{k}_1}_{=0} + \cdots + \frac{\beta_s}{\beta_1}\underbrace{\mathbf{k}_s^{\mathsf{T}}\mathbf{k}_1}_{=0} = \frac{\beta_1}{\beta_1} = 1.$$