Linear programming

Yuval Filmus

November 23, 2017

A *linear program* over the variables $x_1, \ldots, x_n$ is a program of the form

$$\begin{aligned} &\max \sum_{i=1}^n c_i x_i \\ &s.t. \sum_{i=1}^n a_{ji} x_i \leq b_j \text{ for } 1 \leq j \leq m \end{aligned}$$

We can also minimize instead of maximize. The function $\sum_{i=1}^n c_i x_i$ is known as the *objective function*, and the m inequalities are known as *constraints*.

A vector $\vec{x} = x_1, \ldots, x_n$ is *feasible* if it satisfies all constraints. A linear program is *feasible* if there exists a feasible vector. A constraint $\sum_i a_{ji} x_i \leq b_j$ is *tight* for $\vec{x}$ if $\sum_i a_{ji} x_i = b_j$. The *value* of a vector $\vec{x}$ is $\sum_i c_i x_i$. A linear program is *bounded* if its value is bounded over all feasible vectors.

The *coefficient matrix* of the linear program is the $m \times n$ matrix A whose entries are a_{ji}.

1. **Basic feasible solutions.** Consider a linear program which is bounded and feasible.

   (a) Let ϕ be the supremum of the objective function over all feasible vectors. Show that ϕ is achieved at some feasible vector $\vec{x}$. Such vectors are known as *optimal solutions*, and ϕ is the *optimal value* of the program.

   Let $\vec{x}$ be an optimal solution maximizing the number of tight constraints.

   (b) Denote by J the set of tight constraints for $\vec{x}$, and suppose that there exists $1 \leq k \leq m$ such that $\vec{a}_k$ is not in the span of $\{\vec{a}_j : j \in J\}$. Show that there exists a vector $\vec{y}_k$ which is orthogonal to $\{\vec{a}_j : j \in J\}$ but not to $\vec{a}_k$.

   (c) Show that for small enough $\epsilon > 0$, the vectors $\vec{x} \pm \epsilon \vec{y}_k$ are both feasible.

   (d) Deduce that $\vec{y}_k$ is orthogonal to $\vec{c}$.

   (e) By considering the line $\vec{x} + t\vec{y}_k$, show that there is an optimal solution which has more tight constraints than $\vec{x}$.

   (f) Conclude that there exists an optimal solution which satisfies at least rank A linearly independent constraints, where A is the coefficient matrix of the program.

2. **Threshold functions.** A threshold function is a function $\{0,1\}^n \to \{0,1\}$ of the form

$$f(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^n \alpha_i x_i \geq \beta, \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Show that if f is a threshold function then there is an equivalent threshold function for which $|\sum_i \alpha_i x_i - \beta| \geq 1$ for all $\vec{x} \in \{0,1\}^n$.

(b) Consider the following linear program over the variables $\alpha_1, \dots, \alpha_n, \beta$:

$$\begin{aligned} &\max\ 0 \\ &s.t. \sum_{i=1}^{n} \alpha_i x_i \geq \beta + 1 \text{ for all } \vec{x} \in f^{-1}(1), \\ &\sum_{i=1}^{n} \alpha_i x_i \leq \beta - 1 \text{ for all } \vec{x} \in f^{-1}(0). \end{aligned}$$

Show that this program is feasible and bounded.

(c) Conclude that there is a feasible vector $\alpha_1, \dots, \alpha_n, \beta$ which is the solution of a set of $n+1$ linearly independent equations with coefficients in $\{0, \pm 1\}$.

(d) Use Cramer's rule to show that $\alpha_1, \dots, \alpha_n, \beta$ are rational numbers with numerators of magnitude at most $(n+1)!$ and a common denominator of magnitude at most $(n+1)!$.

(e) Conclude that f can be represented using integers $\alpha_1, \dots, \alpha_n, \beta$ of magnitude at most $(n+1)!^2 = e^{O(n \log n)}$.

Håstad showed that the bound $e^{O(n \log n)}$ is tight. It is a nice challenge to give a threshold function which requires exponentially large weights.

3. **Weak duality.** A linear program is in *standard form* if it has the form

$$\begin{aligned} &\max \sum_{i=1}^{n} c_i x_i \\ &s.t. \sum_{i=1}^{n} a_{ji} x_i \leq b_j \text{ for } 1 \leq j \leq m \\ &x_1, \dots, x_n \geq 0 \end{aligned}$$

(a) Show that for every linear program P with n variables and m constraints there is a linear program P' in standard form with the same objective value which has $2n$ variables and the same value of m. (Hint: Write $x = x_+ - x_-$.)

(b) Consider the following two programs:

$$\begin{aligned} &\max \sum_{i=1}^{n} c_i x_i \\ &s.t. \sum_{i=1}^{n} a_{ji} x_i \leq b_j \text{ for } 1 \leq j \leq m \\ &x_1, \dots, x_n \geq 0 \end{aligned} \qquad \begin{aligned} &\min \sum_{j=1}^{n} b_j y_j \\ &s.t. \sum_{j=1}^{m} y_j a_{ji} \geq c_i \text{ for } 1 \leq i \leq n \\ &y_1, \dots, y_m \geq 0 \end{aligned}$$

Show that the value of any feasible solution to the left program (the *primal*) is bounded by the value of any feasible solution to the right program (the *dual*).

Strong duality states that when the primal is feasible and bounded then the dual is also feasible and bounded, and furthermore the optimal values of both coincide.

Optional exercise on following page.

Optional exercise: Yao's minimax principle. The n-dimensional simplex is

$$\Delta_n = \{(x_1, \ldots, x_n) : x_1, \ldots, x_n \geq 0 \text{ and } x_1 + \cdots + x_n = 1\}.$$

Strong duality implies the *minimax principle*: For any coefficients a_{ij},

$$\max_{\vec{x} \in \Delta_m} \min_{\vec{y} \in \Delta_n} \sum_{i=1}^{m} \sum_{j=1}^{n} x_i a_{ij} y_j = \min_{\vec{y} \in \Delta_n} \max_{\vec{x} \in \Delta_m} \sum_{i=1}^{m} \sum_{j=1}^{n} x_i a_{ij} y_j.$$

Let $f \colon \{0,1\}^n \to \{0,1\}$. We consider randomized algorithms for computing f (with some error) of the following form:

- Draw a coordinate i_1 according to some distribution, and query the value of the i_1th input z_{i_1}.
- Draw a coordinate i_2 according to some distribution depending on z_{i_1}, and query the value of z_{i_2}.
- Continue this way until t coordinates $i_1, \ldots, i_t$ are queried.
- Guess the output using a distribution which depends on $z_{i_1}, \ldots, z_{i_t}$.

An algorithm is called *deterministic* if all the distributions encountered in the algorithm are deterministic (one of the possibilities has probability 1).

(a) Show that there is a finite number of deterministic algorithms.

(b) Show that if A is a randomized algorithm then there is a probability distribution $y(A)$ over deterministic algorithms such that A can be simulated by drawing a deterministic algorithm according to $y(A)$ and executing it.

Denote by N the number of deterministic algorithms and by $M = 2^n$ the number of inputs. Let a_{ij} be 1 if the jth deterministic algorithm is wrong on the ith input, and 0 if it is correct.

(c) Show that $\sum_j a_{ij} y(A)_j$ is the error probability of A on the input i.

(d) Deduce that $\max_{\vec{x} \in \Delta_M} \sum_{ij} x_i a_{ij} y(A)_j$ is the error probability of A on the worst input.

(e) Show that $\sum_i x_i a_{ij}$ is the probability that the jth deterministic algorithm makes an error if the input is chosen according to $\vec{x}$.

(f) Deduce that $\min_{\vec{y} \in \Delta_N} \sum_{ij} x_i a_{ij} y_j$ is the error probability of the best deterministic algorithm on the input distribution $\vec{x}$.

(g) Conclude that if there exists a distribution with respect to which every deterministic algorithm has error at least ϵ, then every randomized algorithm makes an error of at least ϵ on some input.