David Bartl

An application of semi-infinite linear programming: approximation of a continuous function by a polynomial

Acta Mathematica Universitatis Ostraviensis, Vol. 12 (2004), No. 1, 3--11

Persistent URL: http://dml.cz/dmlcz/120598

An Application of Semi-Infinite Linear Programming: Approximation of a Continuous Function by a Polynomial

David Bartl

Abstract. We investigate the problem of approximation of a continuous function on a bounded closed interval by a polynomial. We utilise the theory of (semi)-infinite linear programming when treating the problem. At the end of this paper (in Appendix), the utilised Duality Theorem for infinite linear programming is proved.

1. The problem

The problem of approximation (in the sense of minimization of the maximal error) of a continuous function by polynomial can be formulated as follows: Let n be a natural number and let a continuous function $f(x)$ defined on an interval $I = \langle a, b\rangle$ be given. Find a polynomial $P(x)$ of degree $\leq n$ to minimize the maximal error $\varepsilon = \max_{x\in I}\bigl|f(x) - P(x)\bigr|$.

We are especially interested in approximation of the function $f(x) = 1/(1-x)$ on the interval $I = \bigl\langle 0, \frac{1}{2}\bigr\rangle$. We would like to find the best approximative polynomial for this function on the given interval I. (Should this be too difficult, it suffices to find the minimal value of the error $\varepsilon = \max_{x\in I}\bigl|f(x) - P(x)\bigr|$ where $P(x)$ runs through the space of all polynomials with real coefficients of degree $\leq n$. Or find some upper and lower bound for the optimal value of ε at least. Cf. [11].)

We note that the stated problem can have practical applications. It is widely known that mathematical co-processors (found in every PC) use approximative polynomials to compute the value of most mathematical functions (like $\sin x$, $\sqrt{x}$, e^x, as well as x^{-1}, etc.). The operation of division $a \div b$ can be implemented as multiplication by the reciprocal value, $a \times b^{-1}$. The computer represents the "real" number b in the form $m\cdot 2^e$ where $m \in \langle 1, 2)$ is the mantissa and the integer number

Received: July 7, 2004.
2000 *Mathematics Subject Classification:* 41A10, 41A50, 41A52, 90C05, 90C34, 90C90.
Key words and phrases: Approximation of a Continuous Function by a Polynomial, Best Approximation, Semi-Infinite Linear Programming, Infinite Linear Programming, Duality Theorem.

e is the exponent. Then $b^{-1} = 2m^{-1} \cdot 2^{-e-1}$. But $1/2m^{-1} = 1 - x \in \left\langle \frac{1}{2}, 1\right)$ for a suitable $x \in \left(0, \frac{1}{2}\right\rangle$.

2. Preliminaries and notation

Let $f(x)$ always denote the function $f(x) = 1/(1-x)$ and let I always denote the interval $I = \left\langle 0, \frac{1}{2}\right\rangle$ till the end of this paper. Finally, when speaking of a polynomial $P(x)$, we shall always mean some polynomial with real coefficients of degree $\leq n$ where n is a given natural number, also fixed till the end of this paper. The purpose of this paper is to try to solve the above stated problem. But we note the results obtained here are valid, e.g., for any real continuous function f which is defined on a bounded closed interval I and such that for every polynomial $P(x)$ of degree $\leq n$, the function $\varphi(x) = f(x) - P(x)$ does not attain more than n local extremes inside the interval I.

If $I = \left\langle 0, \frac{1}{2}\right\rangle$ is a compact interval, we denote the space of all real continuous functions defined on the interval I as $\mathcal{C}_I$. We endow the space with the classical max-norm $\|\varphi\| = \max_{x\in I}\left|\varphi(x)\right|$ for any $\varphi \in \mathcal{C}_I$. As $\mathcal{C}_I$ is a Banach space and the polynomials of degree $\leq n$ form its subspace of finite dimension, it is easy to see that for any continuous $f \in \mathcal{C}_I$ there exists a polynomial $P(x)$ of degree $\leq n$ so that the distance $\|f - P\|$ is minimal. Hence the maximal error $\varepsilon = \max_{x\in I}\left|f(x) - P(x)\right|$ attains its minimum value for a suitable polynomial $P(x)$.

Let $P^*(x)$ be (one of) the best polynomial(s) and let $\varepsilon^* = \max_{x\in I}\left|f(x) - P^*(x)\right|$ be the optimal (i.e. minimum) value of the error. Using some specific properties (see the second paragraph below) of the function $f(x) = 1/(1-x)$, we infer that the maximal error is attained at exactly $n+2$ points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$ of the interval $I = \left\langle 0, \frac{1}{2}\right\rangle$. (We do not know the points however.) That is, $\left|f(x) - P^*(x)\right| = \varepsilon^*$ if $x = x_k$ for $k = 0, \ldots, n+1$, and $\left|f(x) - P^*(x)\right| < \varepsilon^*$ if $x \in I$ and $x \neq x_k$ for $k = 0, \ldots, n+1$.

We remark that $f(x) = 1/(1-x) = \sum_{k=0}^{\infty} x^k$. It might seem at the first sight that the polynomial $\hat{P}(x) = \sum_{k=0}^{n} x^k$ could approximate the function $f(x)$ well on the given interval $I = \left\langle 0, \frac{1}{2}\right\rangle$. But in fact, $\hat{P}(x)$ is a rather poor approximation. The maximal error $\varepsilon = \max_{x\in I}\left|f(x) - \hat{P}(x)\right|$ is attained at one point $\tilde{x}_0 = \frac{1}{2}$ and the maximal error $\hat{\varepsilon} = f\left(\frac{1}{2}\right) - \hat{P}\left(\frac{1}{2}\right) = \sum_{k=n+1}^{\infty} 1/2^k = 2^{-n}$ is far from being minimal.

Now we are going to justify that in the optimal case, the maximal error $\varepsilon^* = \max_{x\in I}\left|f(x) - P^*(x)\right|$ is attained at exactly $n+2$ points. It suffices to note that the $(n+1)$-st derivative of $\varphi^*(x) = f(x) - P^*(x)$, where $f(x) = 1/(1-x)$, is positive on $(-\infty, 1) \supseteq I$ since the degree of $P^*(x)$ is $\leq n$. The function $\varphi^*(x)$ has not more than n local extremes there, and $P^*(x)$ can have $n+1$ intersections with $f(x)$ at most there. So the function $\varphi^*(x)$ has not more than $n+2$ local extremes on $I = \left\langle 0, \frac{1}{2}\right\rangle$ — two extremes can be at the end-points of the interval. However, if the maximal error were attained at $k+1 < n+2$ points $0 \leq \tilde{x}_0 < \tilde{x}_1 < \cdots < \tilde{x}_k \leq \frac{1}{2}$ on I, then take the polynomial $\tilde{P}(x)$ of degree $k \leq n$ such that $\tilde{P}(\tilde{x}_i) = f(\tilde{x}_i)$ for $i = 0, \ldots, k$, and consider the convex combination $(1-\lambda)P^*(x) + \lambda\tilde{P}(x)$ for small values of $\lambda > 0$. Then the minimal ε^* could yet be decreased — a contradiction. It follows that the maximal error is attained at exactly $n+2$ points of the interval I.

We can formulate the stated problem as a problem of semi-infinite linear programming: Find the value ε and the coefficients $a_n, \ldots, a_0$ of the polynomial $P(x) = a_n x^n + \cdots + a_1 x + a_0$ to minimize ε s.t. $|f(x) - P(x)| \le \varepsilon$ for all $x \in I = \langle 0, \frac{1}{2} \rangle$. An equivalent formulation is:

$$\begin{aligned} &\text{maximize} \ -\varepsilon \\ &\text{s.t.}\ \text{(A)} \quad a_n x^n + \cdots + a_1 x + a_0 - \varepsilon \le \ \ 1/(1-x)\,, \\ &\qquad \text{(B)} \ -a_n x^n - \cdots - a_1 x - a_0 - \varepsilon \le -1/(1-x)\,, \\ &\qquad\quad \text{satisfy both (A) and (B) for all } x \in \langle 0, \tfrac{1}{2} \rangle\,. \end{aligned} \tag{1}$$

We can see that the problem has two sets of constraints, (A) and (B), and that the interval $I = \langle 0, \frac{1}{2} \rangle$ plays the rôle of an index set. If we pick up one of the constraints, we can say that it is formulated for a certain point $x \in I$. Two constraints are formulated for every $x \in I$. Given a constraint formulated for a certain $x \in I$, we shall say that it is of type (A) if and only if it belongs to the set (A), and that the constraint is of type (B) if and only if it belongs to the set (B).

It will be comfortable if we use more compact notation: We shall denote the objective function of problem (1) as $\boldsymbol{\gamma y}$ where $\boldsymbol{\gamma} = (0 \ \cdots \ 0 \ -1)$ is an $(n+2)$-component row vector and $\boldsymbol{y} = (a_n \ \cdots \ a_0 \ \varepsilon)^T$ is a column vector of variables. Further, $\boldsymbol{Ay} \le \boldsymbol{b}$ will be the common notation for all constraints (of both types (A) and (B)) of problem (1).

We already know that (despite the infinite number of constraints) problem (1) does have an optimal solution $P^*(x)$, ε^*. We have already showed that the maximal error ε^* is attained at exactly $n+2$ points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$. It is equivalent to say that the respective constraints (either of type (A) or (B) formulated for that points) are active at the optimum. Let us denote the $n+2$ active constraints as $\boldsymbol{A}_{\text{act}} \boldsymbol{y} \le \boldsymbol{b}_{\text{act}}$.

Let us also determine which of the active constraints are of type (A) and which ones are of type (B). We know that the points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$ are precisely those at which the function $\varphi^*(x) = f(x) - P^*(x)$ attains its (local, but also global) extremes on $I = \langle 0, \frac{1}{2} \rangle$. Obviously, if there is a maximum at a point x_k, say, then there must be minima at points $x_{k\pm1}$. Equivalently, if the active constraint formulated for the point x_k is of type (B), say, then the active constraints formulated for the points $x_{k\pm1}$ must be of type (A).

We also know that $P^*(x)$ and $f(x) = 1/(1-x)$ can have $n+1$ intersections on $(-\infty, 1)$ at most for the $(n+1)$-st derivative of $\varphi^*(x) = f(x) - P^*(x)$ is positive there. On the other hand, they intersect on every (x_k, x_{k+1}) for $k = 0, \ldots, n$. It follows that $P^*(x_{n+1}) < f(x_{n+1})$ because $f(x) \to +\infty$ as $x \to 1^-$, but $P^*(x)$ is bounded. (Therefore, if $P^*(x_{n+1}) > f(x_{n+1})$, then $P^*(x)$ and $f(x)$ would intersect once again on $(x_{n+1}, 1)$.) Hence, the active constraint formulated for the point x_{n+1} is of type (B), the active constraint formulated for the point x_n is of type (A), the active constraint formulated for the point x_{n-1} is of type (B), etc.

3. Applying the duality theory

Theorem 1. *Let X be a real vector space, let $\alpha_j \colon X \to \mathbb{R}$ and $\gamma \colon X \to \mathbb{R}$ be linear functionals and let b_j be real numbers for $j \in J$ where J is an index set. Let the linear programming problem* max γx s.t. *$\alpha_j x \le b_j$ for all $j \in J$ possess an optimal solution $x^* \in X$.*

Using the Duality Theorem in the backward direction, we conclude that the optimal value of the primal problem max $\boldsymbol{\gamma y}$ s.t. $\boldsymbol{A}_{act}\boldsymbol{y} \leq \boldsymbol{b}_{act}$ is also equal to $-\varepsilon^*$. As $P^*(x)$, ε^* is the optimal solution of the original problem (1), it also satisfies the constraints $\boldsymbol{A}_{act}\boldsymbol{y} \leq \boldsymbol{b}_{act}$ of the last primal problem because many constraints are dropped. In addition, the optimal value did not change. So $P^*(x)$, ε^* must be an optimal solution of the last primal problem as well. But the constraints $\boldsymbol{A}_{act}\boldsymbol{y} \leq \boldsymbol{b}_{act}$ are active at this optimum. If we write them as equalities, then $P^*(x)$, ε^* is the unique solution — because the $(n+2) \times (n+2)$ matrix $\boldsymbol{A}_{act}$ is non-singular — of the problem

$$\begin{aligned}&\text{maximize} \quad -\varepsilon \\ &\text{s.t.} \quad \sigma_k \cdot a_n x_k^n + \cdots + \sigma_k \cdot a_1 x_k + \sigma_k \cdot a_0 - \varepsilon = \sigma_k \cdot 1/(1-x_k) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{for } k = 0, \ldots, n+1\end{aligned} \tag{4}$$

where $\sigma_k = (-1)^{n-k}$ for $k = 0, \ldots, n+1$ and the points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$ have the same meaning as above. Hence the surprise: We have managed to reduce the original semi-infinite problem (1) to the problem of solving a system of $n+2$ equations with $n+2$ unknowns $a_n, \ldots, a_0, \varepsilon$. The only problem is that we do not know the points $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$.

4. A try to solve the problem

Let us choose the points $0 = \tilde{x}_0 < \tilde{x}_1 < \cdots < \tilde{x}_n < \tilde{x}_{n+1} = \frac{1}{2}$ arbitrarily, substitute them into (4) and solve the system. We assert that the obtained solution $\tilde{P}(x)$, $\tilde{\varepsilon}$ satisfies $\tilde{\varepsilon} \leq \varepsilon^*$ where ε^* is the true optimal value of the problem. (Should $\tilde{\varepsilon} > \varepsilon^*$ hold, the functions $\tilde{\varphi}(x) = f(x) - \tilde{P}(x)$ and $\varphi^*(x) = f(x) - P^*(x)$ would intersect at $n+1$ points at least, and so would the polynomials $\tilde{P}(x)$ and $P^*(x)$ of degree $\leq n$ — a contradiction.)

Obviously, if the solution $\tilde{P}(x)$, $\tilde{\varepsilon}$ satisfies $\tilde{\varepsilon} < \varepsilon^*$, then it can not be a feasible solution of the original problem (1) — the optimal value ε^* would be lower.

Even the following holds: If $\tilde{\varepsilon} = \varepsilon^*$, then the polynomials $\tilde{P}(x)$ and $P^*(x)$ are the same. (The proof is a rather lengthy exercise and we omit it for that. The main idea is that if the polynomials were different, then the functions $\tilde{\varphi}(x) = f(x) - \tilde{P}(x)$ and $\varphi^*(x) = f(x) - P^*(x)$ — hence both polynomials of degree $\leq n$ — would intersect at $n+1$ points at least.) Therefore, problem (1) possesses exactly one optimal solution.

It follows that we can solve the given problem if, from (4), we express ε in terms of $x_0, \ldots, x_{n+1}$ and minimize it subject to the condition $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = \frac{1}{2}$. The partial derivatives (with respect to $x_1, \ldots, x_n$) must be zero at the extreme. If we knew the point $\boldsymbol{x} = [x_1, \ldots, x_n] \in \mathbb{R}^n$ at which the extreme is attained, we could substitute it back into system (4), learn the coefficients $a_n, \ldots, a_0$ of the optimal polynomial $P^*(x)$, and learn the optimal value of ε again. Although this is a finite-dimensional problem, it is non-linear, and it does not seem to offer an easy solution...

5. Acknowledgement

I would like to thank to Dr. Alan Wilson from London for sending me this interesting problem and asking me for trying to solve it.

Appendix

We shall prove Duality Theorem 2 here. If X is a real vector space, then $X^{\#}$ denotes its algebraic dual, i.e. the space of all linear functionals on X. We endow the algebraic dual with the weak* topology $\sigma(X^{\#}, X)$. (The concept of weak topologies can be found in textbooks on functional analysis. See, e.g., [8, Sections 3, 15, 16].)

We start with a lemma that generalises Farkas' Lemma [5], which is widely known.

Lemma 3 (Farkas' Lemma). *Let X be a real vector space and, J being an index set, let $\alpha_j : X \to \mathbb{R}$ and $\gamma : X \to \mathbb{R}$ be linear functionals on X. Then the implication*

$$\forall j \in J : \ \alpha_j x \leq 0 \implies \gamma x \leq 0$$

holds for all $x \in X$ if and only if

$$\gamma \in \overline{\text{cone}}^* \{ \alpha_j : j \in J \}$$

where $\overline{\text{cone}}^*$ denotes the weakly closed (convex) conical hull of the given subset of $X^{\#}$.*

We do not prove the lemma here because it is a simple consequence of Mazur's Small Theorem on separation of a point from a closed convex set by a hyperplane. The proof of Mazur's Small Theorem can be found in a textbook on functional analysis, e.g., [8, Theorem 14.27]. An elementary proof of Mazur's Small Theorem for the special case of weak topology (which is sufficient here) can be found in [7, Lemma 1]. We note that lemmas very similar to Farkas' Lemma 3, which has just been given above, were formulated a long time ago, see [1, Lemma 1] or [7, Lemma 2]; the authors just did not mention the relationship between their lemmas and Farkas' Lemma.

Before we proceed, let us mention the following fact. If $b \in \mathbb{R}$ is a real number, then it can be considered as a linear functional on $\mathbb{R}$, i.e. as an element of $\mathbb{R}^{\#}$, which is the algebraic dual of $\mathbb{R}$. Indeed, we have the liner mapping $t \mapsto tb$ defined for every $t \in \mathbb{R}$. Consequently, if $\alpha : X \to \mathbb{R}$ is a linear functional on a real vector space X and an inequality $\alpha x \leq b$ is given, then we can interpret it as $\alpha x \leq 1b$ — the linear functional b being evaluated at the point 1 (one). Let us proceed with a lemma which is sometimes considered as a variant of Farkas' Lemma.

Lemma 4 (Lemma on Basic Duality). *Let X be a real vector space. Further, let for $j \in J$, where J is an index set, $\alpha_j : X \to \mathbb{R}$ be linear functionals on X and let $b_j \in \mathbb{R}$ be real numbers (or let $b_j : \mathbb{R} \to \mathbb{R}$ be linear functionals on $\mathbb{R}$). Then the system of inequalities*

$$\alpha_j x \leq b_j \qquad \text{for } j \in J$$

has no solution if and only if

$$(\, o \ \ -1\,) \in \overline{\text{cone}}^* \{ (\, \alpha_j \ \ b_j\,) : j \in J \}$$

where $\overline{\text{cone}}^*$ denotes the weakly closed (convex) conical hull of the given subset of $X^{\#} \times \mathbb{R}^{\#}$.*

Proof. The system of inequalities $\alpha_j x \le b_j$ for $j \in J$ has no solution if and only if the implication

$$\forall j \in J\colon\ \alpha_j x - tb_j \le 0 \implies t = ox + t \cdot 1 = (\,o \quad 1\,)\begin{pmatrix} x \\ t \end{pmatrix} \le 0$$

holds for all $x \in X$ and for all $t \in \mathbb{R}$ (should $t > 0$ held, then x/t would be a solution). Farkas' Lemma 3 finishes the proof. □

Now we give the Weak Duality Theorem, which we will need in the proof of Duality Theorem 2.

Proposition 5 (Weak Duality Theorem). *Let X be a real vector space and let J be an index set. Further, let $\alpha_j\colon X \to \mathbb{R}$ for $j \in J$ and $\gamma\colon X \to \mathbb{R}$ be linear functionals on X, and let $b_j \in \mathbb{R}$ for $j \in J$ be real numbers (or let $b_j\colon \mathbb{R} \to \mathbb{R}$ be linear functionals on $\mathbb{R}$).*

If $x \in X$ satisfies $\alpha_j x \le b_j$ for all $j \in J$ (hence, it is a feasible solution of primal problem (P) *which was stated in Duality Theorem 2 above) and $(\gamma \;\; z) \in \overline{\text{cone}}^*\{\,(\alpha_j \;\; b_j) : j \in J\,\}$, then*

$$\gamma x \le z\,.$$

Proof. If $(\gamma \;\; z) \in \overline{\text{cone}}^*\{\,(\alpha_j \;\; b_j) : j \in J\,\}$, then (by Farkas' Lemma 3) the implication

$$\forall j \in J\colon\ \alpha_j x + tb_j \le 0 \implies \gamma x + tz \le 0$$

holds for all $x \in X$ and for all $t \in \mathbb{R}$. Choose $t = -1$. It follows $\gamma x \le z$, which finishes the proof. □

We can prove Duality Theorem 2 now. The proof of Part I uses an idea which is known from the theory of the classical finite-dimensional linear programming, see [6, "Lemma 4"]. In fact, Part I of Duality Theorem 2 which is stated here further generalises Haar's generalisation of Farkas' Lemma ([9, Theorem 6.1], [2], see also [3, Theorem 4], [4], [10, § 4.II]).

Proof of Duality Theorem 2. I. If z^* is the finite supremum of problem (P), then the implication

$$\forall j \in J\colon\ \alpha_j x \le b_j \implies \gamma x \le z^*$$

holds for all $x \in X$. But we shall see that even

$$\forall j \in J\colon\ \alpha_j x - tb_j \le 0 \implies \gamma x - tz^* \le 0$$

holds for all $x \in X$ and for all $t \in \mathbb{R}$. If this is true, then Farkas' Lemma 3 will finish the proof of Part I. So it remains to prove the implication for all $x \in X$ and $t \in \mathbb{R}$. We distinguish three cases.

If $t > 0$, then the implication is obvious.

Assume that $t = 0$ now. Since $z^* > -\infty$, there exists an $x \in X$ such that $\alpha_j x \le b_j$ for all $j \in J$. If $\alpha_j \hat{x} \le 0$ and $\gamma \hat{x} > 0$ held for some $\hat{x} \in X$ and for all $j \in J$, then $\alpha_j(x + \lambda\hat{x}) \le b_j$ and $\gamma(x + \lambda\hat{x}) > z^*$ for all $j \in J$ and for some large $\lambda > 0$ — a contradiction.

It remains to show the implication for $t < 0$. We can assume without loss of generality that $t = -1$. Assume that $\alpha_j \tilde{x} \le -b_j$ and $\gamma \tilde{x} > -z^*$ held for some $\tilde{x} \in X$ and for all $j \in J$. Then there would exist an $\varepsilon > 0$ such that $\gamma \tilde{x} > -(z^* - \varepsilon)$. Since

z^* is the supremum of the primal problem (P), there would exist an $x^*_\varepsilon \in X$ such that $\alpha_j x^*_\varepsilon \le b_j$ for all $j \in J$ and $\gamma x^*_\varepsilon > z^* - \varepsilon$. Hence $\alpha_j(\tilde{x} + x^*_\varepsilon) \le 0$ for all $j \in J$ and $\gamma(\tilde{x} + x^*_\varepsilon) > 0$. But we already know from above (the case $t = 0$) that this cannot happen.

Part I is proved thus.

II. Note first that the primal problem (P) is feasible. Indeed, if the system $\alpha_j x \le b_j$ for $j \in J$ had no solution, then $(o\ \ -1) \in \overline{\mathrm{cone}}^*\{\,(\alpha_j\ \ b_j) : j \in J\,\}$ by Lemma on Basic Duality 4. Consequently, $(\gamma\ \ z^* - t) \in \overline{\mathrm{cone}}^*\{\,(\alpha_j\ \ b_j) : j \in J\,\}$ for any $t \ge 0$ and z^* could not be the minimum of the dual problem (D) therefore.

Let $\tilde{z}$ be the supremum of the primal problem (P). As that is feasible, we have $\tilde{z} > -\infty$. By using the Weak Duality Theorem 5, we have $\tilde{z} \le z^*$, so the supremum is finite. We can use Part I, which is already proved. We conclude $\tilde{z} = z^*$, which finishes the proof of the theorem. $\square$

References

[1] DUFFIN, R. J. Infinite Programs. In *Linear Inequalities and Related Systems.* Ed. H. W. KUHN, A. W. TUCKER. Princeton: Princeton Univ. Press, 1956, pp. 157–170. (Annals of Mathematics Studies; no. 38.)

[2] HAAR, A. Über lineare Ungleichungen. *Acta Sci. Math.*, 1924, vol. 2, pp. 1–14.

[3] FAN, K. On Systems of Linear Inequalities. In *Linear Inequalities and Related Systems.* Ed. H. W. KUHN, A. W. TUCKER. Princeton: Princeton Univ. Press, 1956, pp. 99–156. (Annals of Mathematics Studies; no. 38.)

[4] FAN, K. – GLICKSBERG, I. – HOFFMAN, A. J. Systems of Inequalities Involving Convex Functions. *Proceedings of the American Mathematical Society,* June 1957, vol. 8, pp. 617–622.

[5] FARKAS, J. Theorie der einfachen Ungleichungen. *Journal für die reine und angewandte Mathematik,* 1902, vol. 124, pp. 1–27.

[6] GRYGAROVÁ, L. *Úvod do lineárního programování.* [*An Introduction to Linear Programming,* in Czech.] Praha: SPN, 1975.

[7] CRAVEN, B. D. – KOLIHA, J. J. Generalizations of Farkas' Theorem. *SIAM Journal on Mathematical Analysis,* November 1977, vol. 8, no. 6, pp. 983–997.

[8] LUKEŠ, J. *Zápisky z funkcionální analýzy.* [*Manuscripts on Functional Analysis,* in Czech]. 1. vydání. [1st edidion.] Praha: Karolinum, 1998. 2. vydání. [2nd edidion.] Praha: Karolinum, 2002.

[9] PRÉKOPA, A. On the Development of Optimization Theory. *American Mathematical Monthly,* August–September 1980, vol. 87, pp. 527–542.

[10] WEYL, H. Elementare Theorie der konvexen Polyeder. *Commentarii Mathematici Helvetici,* 1935, vol. 7, pp. 290–306. English translation: The elementary theory of convex polyhedra. In *Contributions to the Theory of Games. Volume I.* Ed. H. W. KUHN, A. W. TUCKER. Princeton: Princeton Univ. Press, 1950, pp. 3–18. (Annals of Mathematics Studies; no. 24.)

[11] Problem no. A.318. *KöMaL,* April 2003, vol. 53, no. 4, p. 255.

Author(s) Address(es):

DEPT. OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY OF OSTRAVA, 30. DUBNA 22, 701 03 OSTRAVA, CZECH REPUBLIC

E-mail address: bartl@osu.cz