

Multiplicative algorithms for finding nonnegative solution of convex programming problems

by

Yu. S. Popkov

Institute for System Analysis,
Russian Academy of Sciences
9, Prospekt 60 Let Oktyabrya,
Moscow, 117312

An iterative algorithm based on multiplicative transformations is proposed for finding saddle-points of Lagrangians for convex programs. The algorithm uses complementarity conditions, instead of the gradient of the Lagrangian. The convergence of multiplicative algorithms is established by using differential approximations of difference equations and the Lyapunov function of a special type. The multiplicative algorithm is used for solving some issues in urban planning, transportation research and image reconstruction that can be described by the problem of entropy maximization. Results of computational experiments are presented.

1. Introduction

Numerical methods for solving convex programming problems (CPP) are well developed (see, for example, Polyak, 1987). Nevertheless, they can still be improved.

In this paper we investigate a new method for seeking nonnegative solutions of CPP's proposed in Dubov, Imelbayev, Popkov (1983). The method uses the multiplicative scheme for constructing iteration process, namely

$$x^{s+1} = B(x^s, \gamma) = x^s \otimes p(x^s, \gamma),$$

where: γ is a vector of control parameters, x^s is the current iterate and \otimes refers to coordinate-wise multiplication.

This scheme is applied to the search for the saddle-point of the Lagrange function in CPPs. Therefore the numerical method used for this is classified as dual.

The algorithms for finding the Lagrange function saddle-point are well-known (for example, the algorithm of Arrow-Hurwitz-Uzawa (AHU) (Arrow,

Hurwitz, Uzawa, 1958), and the algorithm of augmented Lagrange function, Polyak (1987).

The method proposed differs from the previous ones in the following:

- (i) The multiplicative scheme of iteration process is used which has weaker convergence conditions than in the methods mentioned.
- (ii) The complementarity conditions are used instead of the components of the Lagrange function gradient.
- (iii) The minimum dimension of the state space is realized (no projections on the positive orthant with respect to primary variables are used).
- (iv) One control parameter is used.

In this paper we obtain the convergence conditions for multiplicative algorithms. For these purposes differential approximations of corresponding difference equations and the Lyapunov function of a special form are used.

Computational experiments were carried out on the test problem of entropy maximization on a convex set described by linear inequalities. This is used as a model in many applied problems of urban and regional planning (Wilson, 1970; Popkov et al., 1983), distribution of transportation flows (Shmulyan, Imelbayev, 1978) and image recovering by projections (Gull, Skilling, 1984; Herman 1982; Myrheim, Rue, 1992).

2. Problem statement

Let us consider the convex programming problem

$$\min f(x), g_i(x) \geq 0, i \in \overline{1, r}; x \in R_+^m, m < r, \quad (1)$$

where the functions $f(x)$ and $g_i(x)$ ($i \in \overline{1, r}$) are twice continuously differentiable on R^m ; $f(x)$ is strictly convex and its Hessian is positive defined at x^* (the unique solution of (1)), $g_i(x)$ ($i \in \overline{1, r}$) are concave functions and $R_+^m = \{x : x_i \geq 0, i \in \overline{1, m}; x \in R^m\}$.

Without loss of generality we can assume that

$$\begin{aligned} g_1(x^*) = \dots = g_l(x^*) = 0; x_1^* = \dots = x_k^* = 0; \\ g_{l+1}(x^*) > 0, \dots, g_r(x^*) > 0; x_{k+1}^* > 0, \dots, x_m^* > 0. \end{aligned} \quad (2)$$

We require that the gradients of active constraints be linearly independent at l^* (regularity conditions, see Polyak, 1987) and introduce the Lagrange function for the problem (1):

$$L(x, \lambda) = P(x, \lambda) - \sum_{j=1}^m \lambda_{r+j} x_j, \quad (3)$$

where $P(x, \lambda)$ is the Lagrange function for the problem (1) without nonnegativity conditions:

$$P(x, \lambda) = f(x) - \sum_{i=1}^r \lambda_i g_i(x), \quad (4)$$

$$x \in R_+^m, \lambda \in R_+^{r+m}.$$

Assume that strict complementarity holds, i.e., there exists a pair λ^* such that

$$\lambda_{l+1}^* = \dots = \lambda_r^* = \lambda_{r+k+1}^* = \dots = \lambda_{r+M}^* = 0; \quad (5)$$

$$\lambda_1^* > 0, \dots, \lambda_l^* > 0; \lambda_{R+1}^* > 0, \dots, \lambda_{r+k}^* > 0. \quad (6)$$

Note that the number of variables in the Lagrange function (3) is $2m + r$.

We will use the optimality conditions expressed in terms of $P(x, \lambda)$. These are:

$$P_{x_j}(x^*, \lambda^*) \geq 0; x_j^* P_{x_j}(x^*, \lambda^*) = 0; x_j^* \geq 0; j \in \overline{1, m}; \quad (7)$$

$$P_{\lambda_i}(x^*, \lambda^*) \geq 0; \lambda_i^* P_{\lambda_i}(x^*, \lambda^*) = 0; \lambda_i^* \geq 0; i \in \overline{1, r}.$$

These relations constitute also the existence conditions for the saddle point of the Lagrange function(3). There are $m + r$ variables in (7).

We look for solutions of the problem (1) by means of the following multiplicative algorithm:

$$x_j^{s+1} = x_j^s Q_j(x^s, \lambda^s), j \in \overline{1, m}; \quad (8)$$

$$\lambda_i^{s+1} = \lambda_i^s K_i(\lambda^s, \lambda^s), i \in \overline{1, r};$$

where

$$Q_j(x, \lambda) = 1 - \gamma \frac{\partial P(x, \lambda)}{\partial x_j}, j \in \overline{1, m}; \quad (9)$$

$$K_i(x, \lambda) = 1 - \gamma \frac{\partial P(x, \lambda)}{\partial \lambda_i}, i \in \overline{1, r},$$

γ being positive stepsize.

The multiplicative algorithm (8,9) realizes the search for the Lagrange function saddle-point by using the complementarity conditions in (7).

We will investigate \mathcal{G} -convergence of the multiplicative algorithm. Recall that an algorithm \mathcal{G} -converges if there exist a set $\mathcal{G} \subset R_+^m$ containing the point (x^*, γ^*) and a scalar $a(\mathcal{G}) > 0$ such that for all $\gamma \in (0, a(\mathcal{G}))$ and $(x^0, \gamma^0) \in \mathcal{G}$ the algorithm converges to the solution (x^*, γ^*) of (1), and the convergence is linear in a neighborhood of (x^*, γ^*) .

3. Convergence of the algorithm

Investigation of \mathcal{G} -convergence refers to the study of properties of differential equations (Aliev, Dubov, Izmailov, Popkov, 1985):

$$\dot{x}_j = -x_j P_{x_j}(x, \lambda), j \in \overline{1, m}; \quad (10)$$

$$\dot{\lambda}_i = \lambda_i P_{\lambda_i}(x, \lambda), i \in \overline{1, r};$$

where

$$(x, \lambda) \in R_+^{m+r}.$$

LEMMA 3.1 *The point (x^*, λ^*) is the singular point of the differential equations (10) and it is asymptotically stable for any initial conditions $(x^0, \lambda^0) \in \text{int} R_+^{m+r}$.*

The lemma is proved in the Appendix.

LEMMA 3.2 *The Jacobian of the system (10) is Hurwitz at (x^*, λ^*) .*

The lemma is proved in the Appendix.

THEOREM 3.1 *For any compact set $\mathcal{G} \subset R_+^{m+r}$ algorithm (8,9) \mathcal{G} -converges to the point (x^*, λ^*) which is the solution of problem (1).*

Proof 1. Let us show that algorithm (8,9) is the Euler difference approximation of equations (10). We consider a compact set $\mathcal{G} \subset R_+^{m+r}$, a scalar $\epsilon > 0$ and choose a countable and everywhere dense set of points $\{w_n\}$ in \mathcal{G} . Let t_1 be the time at which the solution of (10) reaches the $\frac{\epsilon}{2}$ -neighborhood of (x^*, λ^*) for the first time when starting from the initial point w_1 (t_1 exists by virtue of Lemma 3.1). Next we choose a neighborhood \mathcal{G}_1 of the point w_1 so that the trajectories starting from the points of this neighborhood will reach the $\frac{3}{4}\epsilon$ -neighborhood of (x^*, λ^*) in time t_1 (this is possible, because the solutions of the differential equation depend continuously on the initial data).

We choose the point w_2 not belonging to \mathcal{G}_1 . Next we determine similarly the neighborhood \mathcal{G}_2 for this point and t_2 . By repeating the process, we obtain the covering of \mathcal{G} by the open sets $\mathcal{G}_1, \mathcal{G}_2, \dots$, from which a finite subcovering $\mathcal{G}_{i_1}, \dots, \mathcal{G}_{i_m}$ can be chosen (because \mathcal{G} is compact). Let $T = \min\{t_{i_1}, \dots, t_{i_m}\}$. Then it follows from the properties of functions f and g_i ($i \in \overline{1, r}$) in problem (1) that we can choose a sufficiently small $\gamma > 0$ such that the trajectories of (1) can be approximated by the Euler piece-wise linear functions within the accuracy $\frac{\epsilon}{4}$ in the segment $[0, T]$ (i.e. the piece-wise linear functions connect iterations (8,9) with accuracy $\frac{\epsilon}{4}$). Note that the algorithm (8,9) is the Euler difference scheme for the differential equations (10). Thus, there exist $\epsilon > 0$, $\gamma(\epsilon) > 0$ and an integer value s_0 such that in (8,9) for all $\gamma \in (0, \gamma(\epsilon))$, $(x^0, \gamma^0) \in \mathcal{G}$ and $s > s_0$

$$\|x^s - x^*\| + \|\gamma^s - \gamma^*\| < \epsilon. \quad (11)$$

2. As functions $f(x)$ and $g_i(x)$ ($i \in \overline{1, r}$) are twice continuously differentiable on R^m then algorithm (8,9) on the set (11) can be described in the following way:

$$u^{s+1} = (E + \gamma J)u^s + \Omega(u^s). \quad (12)$$

where: $u = z - z^*$, $z = \begin{pmatrix} x \\ \gamma \end{pmatrix}$,

$$J = \begin{bmatrix} \frac{\partial}{\partial x_p}(-x_q P_{x_q}) & \frac{\partial}{\partial x_p}(\gamma_s P_{\lambda_s}) \\ \frac{\partial}{\partial \gamma_t}(-x_q P_{x_q}) & \frac{\partial}{\partial \gamma_t}(\gamma_s P_{\lambda_s}) \end{bmatrix}_{(x^*, \gamma^*)}, \quad (13)$$

$p, q \in \overline{1, m}$; $s, t \in \overline{1, r}$; p and t correspond to the rows, and q and s correspond to the columns of J . $\Omega(u^s)$ is a vector function characterizing the remainder of Taylor series.

The iteration process (12) converges to the point (x^*, γ^*) due to Lemma 3.2, see Aliev, Dubov, Izmailov, Popkov (1985). ■

Note that the theorem proved is nonlocal, namely we obtain the convergence conditions for algorithm (8,9) for an arbitrary bounded set of initial approximations.

4. Results of computational experiments

The computational properties of the multiplicative algorithm (8,9) were investigated by means of the test problem of entropy maximization on a polyhedron. The latter is used as a mathematical model in many applied problems, namely in urban planning (Wilson, 1970; Popkov, Posokhin, Gutnov, Schmulyan, 1983), interregional migration (Rogers, Willekens, 1970), co-existence of biological communities (Popkov, 1989), reconstruction of images by projections (Gull, Skilling 1984; Herman 1982; Myrheim, Rue 1992) or design of chemical reactors (Gorban, 1984).

The goal function of this problems is the information entropy characterising the statistical mechanism of forming state variables (statistics of Boltzmann, Fermi-Dirac, or Boze-Einstein). The admissible set is described by a system of linear inequalities. The conventional problem is like

$$H(x) \leftarrow \max, x \in \mathcal{B} \cap \mathcal{D}, \quad (14)$$

where

$$\mathcal{D} = \{x : Tx \leq 1\}; \quad (15)$$

$$\mathcal{B} = \begin{cases} x : x_j \geq 0, & j \in \overline{1, m} \text{ or} \\ x : 0 \leq x_j \leq G, & j \in \overline{1, m}; \end{cases} \quad (16)$$

T is the $(r \times m)$ matrix, $\text{rank } T = r$, $m > r$,

$$t_{kj} \geq 0, k \in \overline{1, r}; j \in \overline{1, m}. \quad (17)$$

We consider three types of entropy functions:

$$H(x) = \begin{cases} -\sum_{n=1}^m x_n \ln \frac{x_n}{ea_n}; \\ -\sum_{n=1}^m x_n \ln \frac{x_n(1-a_n)}{a_n} + (G-x_n) \ln(G-x_n); \\ -\sum_{n=1}^m x_n \ln \frac{x_n}{a_n} - (G-x_n) \ln(G-x_n), \end{cases} \quad (18)$$

where a_n and G are parameters of entropy functions.

It is shown in Popkov (1986) that the problems (14-18) have solutions $x^* \in \text{int}\mathcal{B}$, where \mathcal{B} are the sets (16).

Now, the optimality conditions (7) can be transformed to the following form:

$$P_{x_j}(x^*, \lambda^*) = 0; \quad x_j^* > 0; \quad j \in \overline{1, m}; \quad (19)$$

$$P_{\lambda_i}(x^*, \lambda^*) \leq 0; \quad \lambda_i^* P_{\lambda_i}(x^*, \lambda^*) = 0; \quad \lambda_i^* \geq 0; \quad i \in \overline{1, r}. \quad (20)$$

where

$$P_{x_j}((x^*, \lambda^*)) = -\frac{\partial H}{\partial x_j} + \sum_{i=1}^r \lambda_i t_{ij}, \quad j \in \overline{1, m}. \quad (21)$$

$$P_{\lambda_i}(x^*, \lambda^*) = 1 - \sum_{j=1}^m t_{ij} x_j, \quad i \in \overline{1, r}. \quad (22)$$

After transformations we can reduce the problem (14-18) to the solution of the equations

$$\lambda_k [1 - F_k(\lambda)] = 0, \quad \lambda \in R_+^r, \quad k \in \overline{1, r}; \quad (23)$$

where: λ_j are the Lagrange multipliers for problem (14-16); $F_k(\lambda)$ are determined as

$$F_k(\lambda) = \begin{cases} \sum_{n=1}^m \frac{a_n t_{kn} G_n}{\exp\left(\sum_{j=1}^r t_{jn} \gamma_j\right)}, \\ \sum_{n=1}^m \frac{a_n G_n t_{kn}}{a_n + (1-a_n) \exp\left(\sum_{j=1}^r t_{jn} \lambda_j\right)}, \quad k \in \overline{1, r}; \\ \sum_{n=1}^m \frac{a_n G_n t_{kn}}{\exp\left(\sum_{j=1}^r t_{jn} \lambda_j\right) - a_n}, \end{cases} \quad (24)$$

for the entropy functions (18).

To solve equations (23-24), the multiplicative algorithm (8-9),

$$\lambda_k^{s+1} = \lambda_k^s K_k(\lambda^s), \quad k \in \overline{1, r}, \quad (25)$$

where

$$K_k(\lambda^s) = 1 - \gamma(1 - F_k(\lambda^s)), \quad (26)$$

has been used. We also consider the multiplicative algorithm of coordinate-wise type:

$$\begin{aligned} \lambda_{k1}^{s+1} &= \lambda_{k1}^s(1 - F_{k1}(\lambda^s)), \\ \lambda_k^{s+1} &= \lambda_k^s(1 - \gamma(1 - F_k(\lambda_1^s, \dots, \lambda_{k1-1}^s, \lambda_{k1}^s, \lambda_{k1+1}^s, \dots, \lambda_r^s))), \\ k1 &= \arg \max_k |\lambda_k^s(1 - F_k(\lambda^s))|, \quad k \in \overline{1, r}, \quad k = k1, \end{aligned} \quad (27)$$

which is a modification of algorithm (25).

Multiplicative algorithms were investigated numerically by means of the test problem (14-18) with the following parameters:

$$\begin{aligned} a_n &= \left| \sin \frac{2\pi}{\sqrt{m}} \right|; \quad t_{k1} = \frac{d}{G} \frac{b_{k1}}{c_k}; \quad t_{kn} = \frac{b_{kn}}{c_k}, \quad n \geq 2. \\ b_{nn} &= \tilde{a}_n, \quad b_{kn} = \tilde{a}_n \exp(-v|k - n|), \quad k = n, \quad k \in \overline{1, r}; \quad n \in \overline{1, r}; \\ b_{kn} &= \sqrt{k} \ln a_n, \quad k \in \overline{1, r}; \quad n \in \overline{r+1, m}; \\ \tilde{a}_n &= 1 + a_n; \quad c_k = \sum_{n=1}^m b_{kn}; \end{aligned}$$

The study of the computational properties of algorithms (25-27) was done for the following parameters of the test problem: $d = 30$, $G = 50$, $v = 0.66$, $m = (7/5)r$.

The stopping rule is

$$\text{norm} \leq \text{delt}, \quad (28)$$

where: delt is a positive constant,

$$\text{norm} = \sqrt{\sum_{k=1}^r (\lambda_k^s(1 - F_k(\lambda^s)))^2}. \quad (29)$$

The number of iterations s_0 is a minimal solution of inequality (28-29). The parameter of the set of initial points is

$$\rho_0 = \left(\frac{\sum_{k=1}^r (\lambda_k^0 - \lambda_k^*)^2}{\sum_{k=1}^r (\lambda_k^*)^2} \right)^{\frac{1}{2}}, \quad (30)$$

where: λ_k^0 is k -th coordinate of the initial point, λ_k^* is k -th coordinate of the solution of equations (19-20).

The results of the numerical experiments for the algorithm (25) are shown in Figs. 1-3.

The analysis of the computational experiment showed that:

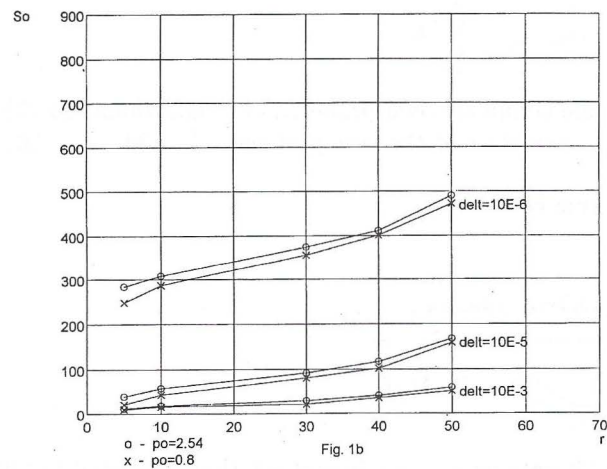
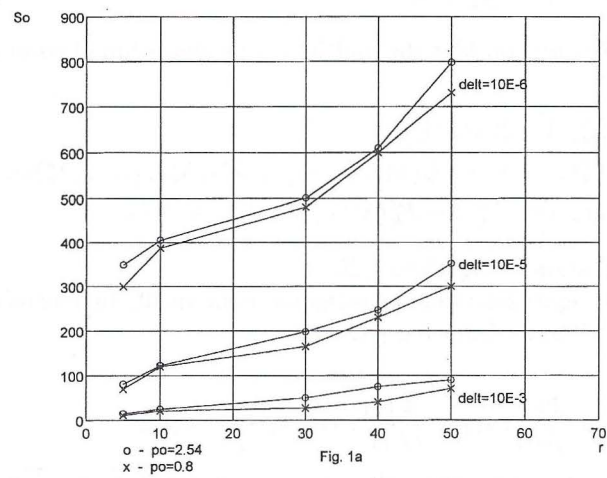


Figure 1. The number of iterations s_0 as a function of dimension r , for different accuracy of calculations ($delt$) and two cases of parameter ρ_0 ; a) algorithm (25); b) algorithm (27).

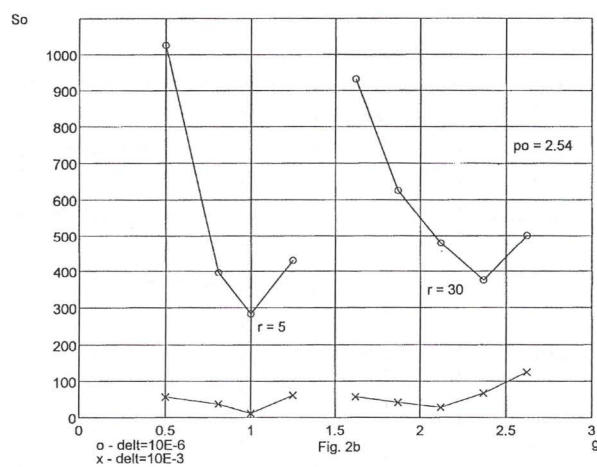
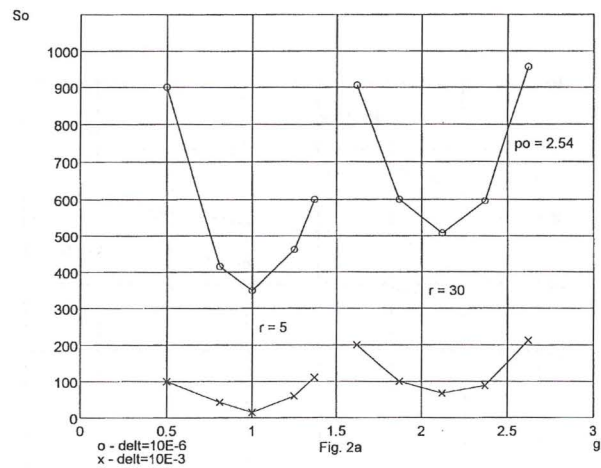


Figure 2. The number of iterations s_0 as a function of stepsize g , for different accuracy (delt) and dimensions r ; a) algorithm (25); b) algorithm (27).

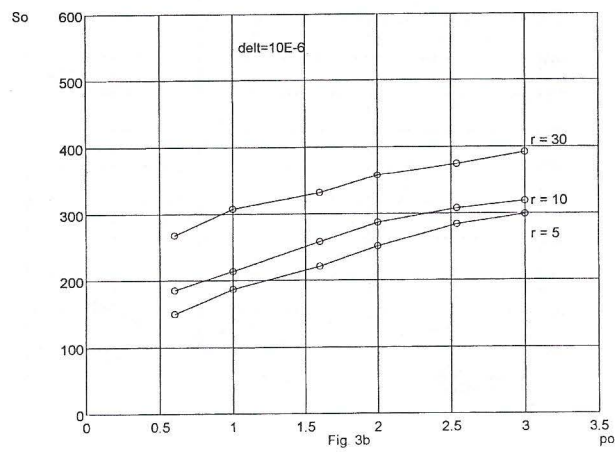
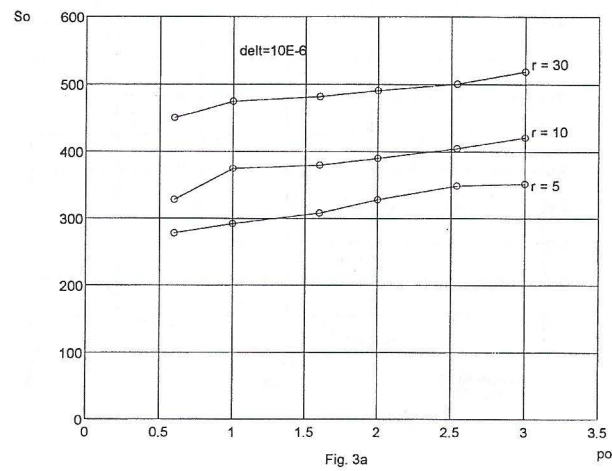


Figure 3. The number of iterations s_0 as a function of ρ_0 , for different dimensions r ; a) algorithm (25); b) algorithm (27).

- (a) the time of computations does not depend essentially on the initial points;
- (b) algorithm (27) converges faster than (25);
- (c) algorithms (27-28) have very fast convergence far from the solution and a very low one in the close neighborhood of it.

References

- ALIEV A.S., DUBOV YU.A., IZMAILOV R.N., POPKOV YU.S. (1985) Convergence of multiplicative algorithm for solving convex programming problems. VNIISI Proceedings *Dynamics of Heterogeneous Systems*, pp. 59-67 (in Russian).
- ARROW K.J., HURWITZ L., UDZAWA H. (1958) *Studies in Linear and Non-linear Programming*, Stanford, Stanford University Press, California.
- DUBOV YU.A., IMELBAYEV SH.S., POPKOV YU.S. (1953) Multiplicative Schemes for Iterative Optimization Algorithms, *Soviet Math. Dokl.*, **28**, 2, pp. 524-526.
- GORBAN A.N. (1984) *Going Around Equilibria*. Novosibirsk, Siberian branch, Nauka (in Russian).
- GULL S.F., SKILLING J. (1984) Maximum entropy method in image processing. *IEE Proc.*, **131**, Pt.F., N6, p. 646-659.
- HERMAN G.T. (1982) Mathematical optimization versus practical performance: a case study based on the maximum entropy criterion in image reconstruction. *Mathem. Progr. Study*, **20**, pp. 96-112.
- IMELBAYEV SH.S., SHMULYAN B.L. (1978) Modeling stochastic communication systems. In: Wilson A.G., *Entropy in Urban and Regional Modeling* (Russian translation). Moscow, Nauka, pp. 230-274 (in Russian).
- KRASNOSELSKY M.A. et. al. (1972) *Approximate Solution of Nonlinear Operator Equations*. Groningen, Nordhoff.
- MYRHEIM J., RUE H. (1992) New Algorithms for Maximum Entropy Image Restoration. *Graphical Models and Image Processing*, **54**, 3, pp. 223-238.
- POLYAK B.T. (1987) *Introduction to Optimization*. New York, Optimization Software.
- POPKOV YU.S., POSOKHIN M.V., GUTNOV A.E., SCHMULYAN B.L. (1983) *Systems Analysis and Urban Development Problems*. Moscow, Nauka (in Russian).
- POPKOV YU. S. (1986) Mathematical models of macrosystem stationary states: principles of formulation, analysis of properties and algorithms. *Large Scale Systems*, **10**, pp. 1-20.
- POPKOV YU. S. (1989) A new class of dynamic macrosystem models with self-reproduction. *Environment and Planning A*, **21**, p. 739-751.
- POPKOV YU. S. (1992) Multiplicative schemes of iteration algorithms. (1. Nonlinear equations). *Technical Cybernetics*, **6**, pp. 110-117.
- ROGERS A., WILLEKENS F. (1978) Migration and Settlement: Measurement and analysis. IIASA Report RR-78-13, August, Laxenburg, Austria.

WILSON A.G. (1970) *Entropy in Urban and Regional Modeling*. Pion, London.

Appendix

1. Proof of lemma 3.1. Let us consider the function

$$\begin{aligned} v(x, \lambda) &= \sum_{j=1}^m (x_j - x_j^*) - x_j^* (\ln x_j - \ln x_j^*) + \\ &+ \sum_{i=1}^r (\lambda_i - \lambda_i^*) - \lambda_i^* (\ln \lambda_i - \ln \lambda_i^*), \end{aligned} \quad (31)$$

$(z^*, \lambda^* \in R_+^{m+r}).$

where $y \ln y = 0$ for $y = 0$.

The function V is positive and convex and attains its minimum $V(x^*, \lambda^*) = 0$. Hence the sets $\mathcal{B}_a = \{(x, \lambda) : V(x, \lambda) \leq a\}$ are bounded, for any bounded set in R_+^{m+r} there exist a scalar $c > 0$ and the set $\mathcal{B}_c = \{(x, \lambda) : V(x, \lambda) \leq c\}$ such that $\mathcal{B}_c \subset \mathcal{B}_a$.

Let us consider the derivate V along the trajectories of (31):

$$\begin{aligned} \frac{dV}{dt} &= \left[f(x) - f(x^*) - \sum_{j=1}^m (x_j - x_j^*) \frac{\partial f(x)}{\partial x_j} \right] + \\ &+ \left[\sum_{i=1}^r \lambda_i \left(\sum_{j=1}^m (x_j - x_j^*) \frac{\partial g_i(x)}{\partial x_j} - (g_i(x) - g_i(x^*)) \right) \right] + \\ &+ \left[\sum_{i=1}^r \lambda_i^* g_i(x) - \sum_{i=1}^r \lambda_i g_i(x^*) - (f(x) - f(x^*)) \right]. \end{aligned} \quad (32)$$

We investigate the sign of (32) on R_+^{m+r} . The second term is nonpositive because $g_i(x)$ ($i \in \overline{1, r}$) is concave; the first term is negative for $x = x^*$ (because $f(x)$ is convex and its Hessian is not degenerated at x^*) and is zero at $x = x^*$. Consider the third term of (32) using the saddle-point definition:

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*). \quad (33)$$

These inequalities should hold for any $\lambda \geq 0$, $x \geq 0$, in particular, for $x \geq 0$, $\lambda_1 \geq 0, \dots, \lambda_{r+k} \geq 0$ and $\lambda_{r+k+1} = \dots = \lambda_{r+m} = 0$. Then we have from the definition of the saddle point:

$$f(x^*) - f(x) - \sum_{i=1}^r \lambda_i g_i(x^*) + \sum_{i=1}^r \lambda_i^* g_i(x) \leq 0,$$

i.e. the third term in (32) is nonpositive.

Therefore the expression in (32) is non-positive in R_+^{m+r} . Moreover, if $x = x^*$ then (32) is negative; if $x = x^*$ then the first and second terms are zero, while, by virtue of (5-6), the third term in (32) is $-\sum_{i=l+1}^r \lambda_i g_i(x^*)$, i.e. is either zero (if $\lambda_{l+1} = \lambda_{l+1}^* = \dots = \lambda_r = \lambda_r^* = 0$) or negative.

Hence

$$\frac{dV}{dt} \begin{cases} < 0, & (x, \lambda) \in R_+^{m+r} \setminus \mathcal{L}, \\ = 0, & (x, \lambda) \in \mathcal{L}, \end{cases} \quad (34)$$

where

$$\mathcal{L} \{ (x, \lambda) : x = x^*; \lambda_{l+1} = \dots = \lambda_r = 0 \}. \quad (35)$$

Note that $(x^*, \lambda^*) \in \mathcal{L}$.

For $l = 0$ we have $\mathcal{L} = (x^*, \lambda^*)$, and V in (31) is the Lyapunov function for (10), namely inside of $\mathcal{B}_c \subset R_+^{m+r}$ and outside of the ϵ -neighborhood of (x^*, λ^*) the derivative of V along the trajectories of system (10) is negative and strictly separated from zero, (34).

Now let $l \geq 1$. Let us consider $P(x, \lambda)$ (4) and determine its derivative along the trajectories of (10). By (5-6), on the set \mathcal{L} , (35), we obtain

$$\frac{dP}{dt} = - \sum_{j=k+1}^m x_j^* \left(\sum_{i=1}^l (\lambda_i - \lambda_i^*) \frac{\partial g_i(x^*)}{\partial x_j} \right)^2. \quad (36)$$

It follows from the linear independence of $\nabla g_1(x^*), \dots, \nabla g_l(x^*)$, e_1, \dots, e_k (regularity condition in problem (1)) that the last $m - k$ (note that $m - k \geq 1$) coordinates of the vectors $\nabla g_1(x^*), \dots, \nabla g_l(x^*)$ are linearly independent. Therefore (37) is a continuous quadratic function of $(\lambda - \lambda^*)$.

Thus

$$\frac{dP}{dt} \begin{cases} < 0, & (x, \lambda) \in \mathcal{L} \setminus (x^*, \lambda^*); \\ = 0, & (x, \lambda) = (x^*, \lambda^*). \end{cases} \quad (37)$$

Denote by \mathcal{L}_δ the δ -neighborhood of the set \mathcal{L} (5), and by \mathcal{L}_ϵ the ϵ -neighborhood of the point (x^*, λ^*) .

Let us consider the function

$$V_\eta(x, \lambda) = V(x, \lambda) + \eta P(x, \lambda), \quad (38)$$

$$(x, \lambda) \in R_+^{m+r}, \eta > 0$$

and the sets

$$\mathcal{B}_c = \{ (x, \lambda) : x, \lambda \leq c \}, \quad (39)$$

$$\mathcal{M}_c(\eta) = \{ (x, \lambda) : V_\eta(x, \lambda) \leq c \}. \quad (40)$$

Since the functions V and P are convex, there exists $\epsilon > 0$ such that

$$\mathcal{B}_c \subset \mathcal{M}_{c+\epsilon}(\eta) \subset \mathcal{B}_{c+2\epsilon}. \quad (41)$$

Since (36) is continuous, there exists $\delta > 0$ such that $\frac{dP}{dt} < 0$ for $(x, \lambda) \in (\text{cal} \mathcal{B}_c \cap \mathcal{L}_\delta) \setminus \mathcal{L}_\epsilon$. According to (41), we can choose $\eta^* > 0$ so that

$$\eta \left| \frac{dP}{dt} \right| \leq \frac{1}{2} \left| \frac{dV}{dt} \right|$$

for all $\eta \in (0, \eta^*)$ and $(x, \lambda) \in \mathcal{I} = (\mathcal{M}_{c+\epsilon}(\eta) \cap \mathcal{L}_\delta) \setminus \mathcal{L}_\epsilon$.

Hence the function $V_\eta(x, \lambda)$ for $(x, \lambda) \in \mathcal{I}$ is the Lyapunov function for system (10), namely its time derivative along the trajectories of (10) is negative and strictly separated from zero; it is zero only at the point (x^*, λ^*) .

Thus, for the system (10) there exists a constant $\eta > 0$ and the function

$$\theta(x, \lambda) = \begin{cases} V(x, \lambda), & (x, \lambda) \in R_+^{m+r} \setminus \mathcal{L}, \\ V(x, \lambda) + \eta P(x, \lambda), & (x, \lambda) \in \mathcal{L}, \end{cases}$$

whose time derivative is negative for all $(x, \lambda) \in R_+^{m+r}$ and zero at (x^*, λ^*) . ■

2. Proof of lemma 3.2. Consider the matrix

$$J = \begin{bmatrix} \frac{\partial}{\partial x_p}(-x_q P_{x_q}) & \frac{\partial}{\partial x_p}(\gamma_s P_{\lambda_s}) \\ \frac{\partial}{\partial \gamma_t}(-x_q P_{x_q}) & \frac{\partial}{\partial \gamma_t}(\gamma_s P_{\lambda_s}) \end{bmatrix}_{(x^*, \gamma^*)}, \quad (42)$$

where $p, q \in \overline{1, m}$; $s, t \in \overline{1, r}$, p and t correspond to the rows, and q and s correspond to the columns of J . Note that, by virtue of (5,6,7), the relations

$$P_{x_j}(x^*, \lambda^*) = L_{x_j}(x^*, \lambda^*) + \lambda_{r+j}^* \geq 0, \quad j \in \overline{1, k};$$

$$P_{x_j}(x^*, \lambda^*) = L_{x_j}(x^*, \lambda^*), \quad j \in \overline{k+1, m}.$$

hold. Let us rewrite (42), taking into account these equations.

$$J = \left[\begin{array}{cc|cc|cc} -\lambda_{r+1}^* & & \cdots & \cdots & & \\ & \ddots & & & & \\ & & -\lambda_{r+k}^* & & & \\ \hline & & & H & U & \\ & 0 & & V & 0 & \\ \hline & & & \cdots & & -g_{l+1} \\ & & & \cdots & 0 & \ddots \\ & & & \cdots & & -g_r \end{array} \right]_{(x^*, \lambda^*)}, \quad (43)$$

where the elements of the matrices H , U and V are equal to

$$h_{pq} = -x_q^* P_{x_p x_q}; \quad u_{ps} = -\lambda_s^* \frac{\partial g_s}{\partial x_p}; \quad v_{tq} = x_q^* \frac{\partial g_t}{\partial x_q}; \quad (44)$$

and

$$p, q \in \overline{k+1, m}; \quad s, t \in \overline{l+1, r},$$

respectively.

The dots in (43) can be replaced by zeros, with J remaining stable. Indeed, the Raus-Hurwitz criterion of matrix stability links the matrix stability with the properties of coefficients in the characteristic polynomial. These coefficients are the sums of main minors. It follows from (43) that none of these minors will be changed in replacing the dots by zeros. Thus, instead of proving the stability of J , we can prove the stability of J_1 , with J_1 differing from J by zeros and an orthogonal transformation of the basis (this does not change the eigenvalues and hence does not influence the stability of J):

$$J = \left[\begin{array}{cccc|cc} -\lambda_{r+1}^* & & & & & \\ & \ddots & & & & \\ & & -\lambda_{r+k}^* & & 0 & \\ & & & -g_{l+1}(x^*) & & \\ & 0 & & & \ddots & \\ & & & & & -g_m(x^*) \\ \hline & & & 0 & & \\ \hline & & & & \frac{H}{V} & \frac{U}{0} \end{array} \right]. \quad (45)$$

The negative-definiteness and nondegeneracy of the matrix $-P_{x_p x_q}(x^*, \lambda^*)$ follows from the convexity of f , concavity of g_i ($i \in \overline{1, r}$) and non-degeneracy of the Hessian of f at X^* . Besides, the matrix $\nabla g(x^*)$ is of full rank (a corollary of the regularity conditions) and the numbers x_q^* , λ_s^* ($q \in \overline{k+1, m}$; $s \in \overline{1, l}$); $\lambda_{r+1}^*, \dots, \lambda_{r+k}^*$; $g_{l+1}(x^*), \dots, g_r(x^*)$ are positive.

The spectrum of J_1 consists of $-\lambda_{r+1}^*, \dots, -\lambda_{r+k}^*, -g_{l+1}(x^*), \dots, -g_m(x^*) < 0$ and eigenvalues of the matrix

$$\tilde{J}_1 = \begin{bmatrix} H & U \\ V & 0 \end{bmatrix},$$

whose elements are given by the equalities (44). ■

Studies in Fuzziness

Series Editor: J. Kacprzyk

M. Delgado, University of Granada, Spain;
J. Kacprzyk, Polish Academy of Sciences, Warsaw,
Poland; **J.-L. Verdegay**, **M.A. Vila**, University of
Granada, Spain (Eds.)

Fuzzy Optimization

Recent Advances

1994. XII, 459 pp. 68 figs., 28 tabs.
(Studies in Fuzziness, Vol. 2)
Hardcover DM 168,- ISBN 3-7908-0749-4

Starting with new concepts of optimality and general approaches to optimization under fuzziness, further issues are related to multiobjective fuzzy optimization and interactive approaches. Relations between fuzzy optimization and such techniques as genetic algorithms, simulated annealing, and neural networks are discussed. Applications to real problems are presented.

A. Geyer-Schulz, Vienna University of Economics
and Business Administration, Austria

Fuzzy Rule-Based Expert Systems and Genetic Machine Learning

1995. XVIII, 413 pp. 67 figs., 16 tabs.
(Studies in Fuzziness, Vol. 3)
Hardcover DM 148,- ISBN 3-7908-0830-X

This book presents fuzzy rule-languages as link between quantitative and qualitative models. In the first part the semantic of fuzzy rule-languages is extended with a type system and an object-oriented system. In the second part fuzzy rule-languages are integrated with genetic algorithms and with classifier systems. For this purpose the class of genetic algorithms over context-free languages has been developed.

T. Onisawa, University of Tsukuba, Japan;
J. Kacprzyk, Polish Academy of Sciences, Warsaw,
Poland (Eds.)

Reliability and Safety Analyses under Fuzziness

1995. X, 376 pp. 106 figs., 48 tabs.
(Studies in Fuzziness, Vol. 4)
Hardcover DM 148,- ISBN 3-7908-0837-7

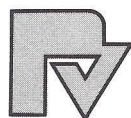
This book provides a comprehensive, up-to-date account on recent applications of fuzzy sets and possibility theory in reliability and safety analysis. It presents new tools for handling non-probabilistic aspects of uncertainty in these problems.

P. Bosc, IRISA/ENSSAT, Lannion, France;
J. Kacprzyk, Polish Academy of Sciences, Warsaw,
Poland (Eds.)

Fuzziness in Database Management Systems

1995. VIII, 433 pp. 100 figs., 32 tabs.
(Studies in Fuzziness, Vol. 5)
Hardcover DM 168,- ISBN 3-7908-0858-X

This book provides a comprehensive treatise on the use of fuzzy sets, fuzzy logic, and possibility theory for dealing with imprecise and vague information in database management systems. Theoretical works and implemented systems for commercial database management systems are discussed. A fuzzy approach to database security is considered.



Physica-Verlag

A Springer-Verlag Company

Please order through your bookseller or from Physica-Verlag,
c/o Springer-Verlag GmbH & Co. KG, Auftragsbearbeitung, Postfach 31 13 40, D-10643 Berlin, Germany