Kybernetika

Tomáš Cipra
Prediction in stochastic linear programming

PREDICTION IN STOCHASTIC LINEAR PROGRAMMING

TOMÁŠ CIPRA

If random values in a linear program with random coefficients can be predicted using previous observations on them, one can utilize this prediction and construct the optimal decision before observing the appropriate coefficients of the program. Various theoretical and practical aspects of this statistical approach to random linear programs are discussed mainly in the situation when we use prediction regions. The results are demonstrated by means of a numerical example.

1. INTRODUCTION

The methods of stochastic programming which can be characterized as the statistical approach start to be applied recently (see e.g. [3], [4], [7]). This trend is comprehensible because it is important to treat correctly and in suitable way from the statistical point of view the data which we have at our disposal for particular stochastic programs. The first phase of the investigation (in which we treat the given observations, calculate the appropriate probability characteristics of the corresponding distributions or make some conjectures on their behaviour) is not less important than the following phase of optimization. The final results depend from the practical point of view not only on the quality of the used optimization procedure but also on the quality of the treatment of the statistical input.

This paper concentrates on the situation in which the observations for a linear program with random coefficients have the form of mutually dependent values, e.g. they form statistical time series. Statistical data of this type are common in many practical disciplines, e.g. in economy, hydrology, biology, social sciences and others because in these cases the dynamic movement in time is mostly important. Moreover, there exists rich statistical methodology including easily acceptable computer software how to treat such data. It is clear that such type of data cannot avoid stochastic programming and it has motivated this contribution.

Let us consider the following linear program

(1.1) $\{\min c'x: Ax = b, \; x \geq 0\}$

the coefficients of which form (multivariate) time series $\{A_t\}, \{b_t\}, \{c_t\}$ of dimensions $(m, n), (m, 1), (n, 1)$, respectively. Better to say, we have finite realizations $A_1, \ldots, A_T$, $b_1, \ldots, b_T$, $c_1, \ldots, c_T$ at our disposal as the data for this program. The statistical dependency of time series values enables to construct predictions of future values (it is one of the basic aims of time series analysis) and therefore the natural approach to the previous program is to work with predicted coefficients. If e.g. the decision for time $T + 1$ is to be constructed at time T before we observe the values at time $T + 1$ we shall construct this decision $\hat{x}_{T+1} = \hat{x}_{T+1}(\hat{A}_{T+1}, \hat{b}_{T+1}, \hat{c}_{T+1})$ from the predicted coefficients $\hat{A}_{T+1}, \hat{b}_{T+1}, \hat{c}_{T+1}$ for time $T + 1$ (or from predictions for more steps ahead).

This procedure can be applied sequentially in time. In sequential treatment (or on-line treatment) we get the data gradually for particular times and at the actual time we exploit all information we have by this time at our disposal. Some tricks of postoptimization are possible to simplify the necessary calculations. If e.g. a periodical pattern with a period d is recognizable in the data we can examine before solving the program (1.1) at time $T + 1$ whether possibly the same position of the basis as the position of the optimal basis at time $T + 1 - d$ is also optimal for time $T + 1$. In the case of a mild linear trend we can try the same trick but for the previous time $(T + 1) - 1 = T$, etc. Practical experiences [5] show that such procedure may be successful even using very simple prediction methods as the simple and double exponential smoothing, Holt-Winters seasonal method and others (see e.g. [2], [13]).

As the prediction is concerned practitioners prefer sometimes so called prediction regions to the above mentioned point predictions ($(1 - \alpha)$ prediction region is the set in which the predicted value lies with probability $1 - \alpha$ where $0 < \alpha < 1$ is a given constant usually near to zero). The reason for this preference is that e.g. for an airline company it is more interesting to find that the number of passengers in the future month will not be lower than 100 000 with 95 per cent certainty than to give some exact prediction number without any probability description.

The remaining parts of the paper suggest how to utilize such prediction regions in the program (1.1). We confine ourselves for simplicity to the case when the right-hand side b of (1.1) is random only (i.e. A and c are deterministic) but extension to the general case with full randomness is possible and is referred to in the text (the case with random c only is covered at once due to duality).

If we have constructed the appropriate prediction region for b (we shall show in Section 2 that such prediction region is frequently an elipsoid in practice) we should investigate at first whether this prediction region is the subset of such region for b in which the program (1.1) is solvable (so called solvability region). This problem is discussed in Section 3 while Section 4 is devoted to various possibilities how to exploit the prediction region for b when solving (1.1). The methods of parametric program-

ming are used in Sections 3 and 4 which enable to obtain exact analytical results. The disadvantage of parametric programming algorithms is that they are elaborate and therefore Section 5 offers application of so called bunching method which provides comparable results and is more convenient for practical calculations. A simple example in Section 6 demonstrates the suggested procedures numerically.

2. CONSTRUCTION OF PREDICTION REGIONS IN PRACTICE

There are various possibilities how to construct the prediction regions in practice. The following two methods which are frequent in practical applications demonstrate that we can confine ourselves to prediction regions in the form of elipsoids:

a) Prediction based on econometric modeling (see e.g. [11, p. 406]) is a general method which includes as a special case e.g. the prediction based on the classical regression analysis. If using this method we must e.g. have at our disposal the estimated model of simultaneous equations in the reduced form

$$(2.1) \qquad y_t = \Pi x_t + v_t, \quad t = 1, \ldots, T.$$

Here $y_t = (y_{1t}, \ldots, y_{mt})'$ is a vector of endogenous variables at time t which is explained by a vector of predetermined variables $x_t = (x_{1t}, \ldots, x_{kt})'$ at time t (the object of the prediction are the endogenous variables), Π is an (m, k) matrix of parameters and $v_t = (v_{1t}, \ldots, v_{mt})'$ is a vector of disturbances at time t. One assumes that v_t are normally distributed with $\mathsf{E}v_t = 0$, $\mathsf{E}v_t v_t' = \Sigma_{vv}$ (a positive-definite matrix) and $\mathsf{E}v_s v_t' = 0$ for $s \neq t$. The model can be summarized for all t as

$$(2.2) \qquad Y = X\Pi' + V,$$

where $Y = (y_1, \ldots, y_T)'$, $X = (x_1, \ldots, x_T)'$ and $V = (v_1, \ldots, v_T)'$. The classical OLS (Ordinary Least Squares) estimator of Π has the form

$$(2.3) \qquad \hat{\Pi} = Y'X(X'X)^{-1}.$$

Let $\hat{x}_{T+h}$ be a vector of predicted predetermined variables for time $T + h$ which can be obtained in some way outside the model, see discussion in [9, p. 196] (e.g. in a simple regression situation it can be $x_{it} = t$ so that $\hat{x}_{i,T+h} = T + h$). Then under general assumptions on the stochastic behavior of the model at time $T + h$ (there must not be a change in the specification of the model at this time) the optimal point prediction for the endogenous variable at time $T + h$ can be constructed as

$$(2.4) \qquad \hat{y}_{T+h} = \hat{\Pi}\hat{x}_{T+h}$$

with the $(1 - \alpha)$ prediction region of the form

$$(2.5) \qquad (y_{T+h} - \hat{y}_{T+h})' S_{T+h,T+h}^{-1}(y_{T+h} - \hat{y}_{T+h}) \leqq$$

$$\leqq \frac{(T - k)\, m}{T - k - m + 1}\, F_{m,T-k-m+1}(\alpha).$$

Here

(2.6) $S_{T+h,T+h} = \dfrac{1}{T-k}\left[1 + \hat{x}'_{T+h}(X'X)^{-1}\,\hat{x}_{T+h}\right]\left(Y'Y - \hat{\Pi}X'Y\right)$

is the estimated convariance matrix of the error $y_{T+h} - \hat{y}_{T+h}$ of the prediction and $F_{m,T-k-m+1}(\alpha)$ is the tabulated critical value of Fisher's distribution with the appropriate degrees of freedom and the level of significance α (e.g. (2.5) is the 95% prediction region for $\alpha = 0{\cdot}05$). So called Hotelling's statistic has been used to derive (2.5). The assumption of normality for v_t can be replaced by more general assumptions under which (2.5) holds asymptotically (see e.g. [6, p. 161]).

Remark 1. Various multivariate trends can be modeled by means of (2.1). E.g. the multivariate polynomial trend is modeled as

(2.7) $y_{it} = f_i(t) + v_{it}\,, \quad i = 1, \ldots, m\,, \quad t = 1, \ldots, T\,,$

where $f_i(t)$ is a polynomial of order p_i (it means that the predetermined variables are chosen as powers of time t).

Remark 2. Hymans [10] suggested $(1 - \alpha)$ joint prediction intervals for particular components of y_{T+h} in the form

(2.8) $\left(\hat{y}_{i,T+h} - \sqrt{(cs_{ii})},\ \hat{y}_{i,T+h} + \sqrt{(cs_{ii})}\right)\,, \quad i = 1, \ldots, m\,,$

where s_{ii} is the ith diagonal element of the matrix $(Y'Y - \hat{\Pi}X'Y)/(T - k)$ and

(2.9) $c = \dfrac{(T-k)\,m}{T-k-m+1}\left[1 + \hat{x}'_{T+h}(X'X)^{-1}\,\hat{x}_{T+h}\right]F_{m,T-k-m+1}(\alpha)\,.$

b) Prediction in the framework of Box-Jenkins approach (see [1], [9]) is exploited by many statisticians and econometricians as a very flexible and fruitful prediction method. Similarly as in the econometric modeling one must construct an appropriate model at first. Box-Jenkins methodology utilizes so called ARMA (p, q) models (or their various modifications) of the form

(2.10) $y_t + A_1 y_{t-1} + \ldots + A_p y_{t-p} = \varepsilon_t + B_1 \varepsilon_{t-1} + \ldots + B_q \varepsilon_{t-q}\,,$

where y_t is the modeled m-dimensional process, $A_1, \ldots, A_p$ and $B_1, \ldots, B_q$ are (m, m) matrices of parameters and ε_t is the m-dimensional normal white noise, i.e. $E\varepsilon_t = 0$, $E\varepsilon_t\varepsilon_t' = \Sigma$ (a positive-definite matrix), $E\varepsilon_s\varepsilon_t' = 0$ for $s \neq t$. Then under general conditions the optimal point prediction for time $T + h$ can be written as

(2.11) $\hat{y}_{T+h} = \displaystyle\sum_{j=0}^{\infty} C_{j+h}\varepsilon_{T-j}\,,$

where the (m, m) matrices C_i fulfill the following power series equation

(2.12) $\left(I + A_1 z + \ldots + A_p z^p\right)\left(I + C_1 z + C_2 z^2 + \ldots\right) = I + B_1 z + \ldots + B_q z^q\,.$

The corresponding $(1 - \alpha)$ prediction region has the form

(2.13) $\left(y_{T+h} - \hat{y}_{T+h}\right)' V(h)^{-1}\left(y_{T+h} - \hat{y}_{T+h}\right) \leqq \chi_m^2(\alpha)\,,$

217

where

$$(2.14) \qquad V(h) = \sum_{j=0}^{h-1} C_j \, \Sigma C_j'$$

is the covariance matrix of the error $y_{T+h} - \hat{y}_{T+h}$ of the prediction and $\chi_m^2(\alpha)$ is the tabulated critical value of chi-squared distribution with m degrees of freedom and the level of significance α.

The following conclusion can be drawn from the previous text. In both prediction methods (and also in other less important ones) the appropriate $(1 - \alpha)$ prediction region has the geometric form of an elipsoid. This elipsoid can be written generally for the program (1.1) as

$$(2.15) \qquad (b - \hat{b})' \, V(b - \hat{b}) \leqq k(\alpha) \, ,$$

where $\hat{b}$ is a known vector (the center of the elipsoid), V is a known positive-definite matrix and $k(\alpha)$ is a known constant. Let us denote this region as

$$(2.16) \qquad P(\alpha) = \{ b \in R^m : (b - \hat{b})' \, V(b - \hat{b}) \leqq k(\alpha) \} \, .$$

3. PROBLEM OF SOLVABILITY

Let us denote the region of solvability of the program (1.1) with the random right-hand side as

$$(3.1) \qquad S = \{ b \in R^m : (1.1) \text{ has an optimal solution} \}$$

(i.e. the program (1.1) is feasible and bounded for all $b \in S$) and assume that S is nonempty. Then in our context the problem of solvability consists in the investigation of the inclusion $P(\alpha) \subset S$.

Wets [18] deals with a general problem of this type when he investigates feasibility of stochastic programs with fixed recourse. Using theory of polar matrices and cone ordering he can treat cases with very general regions $P(\alpha)$. In our case we make use of the special elipsoid shape of the prediction region $P(\alpha)$ and proceed in the following way.

The solvability region S is a convex polyhedral cone with the vertex in the origin (see e.g. [16], [17]), i.e.

$$(3.2) \qquad S = \{ b \in R^m : h_i' b \geqq 0, \ i = 1, \dots, N \} \, .$$

The explicit numerical form of this cone (i.e. the vectors $h_1, \dots, h_N$) can be found by means of various algorithmic procedures (e.g. [8], [14, p. 276]).

In order to simplify the solution of our problem let us transform the coordinate system in R^m so that the elipsoid (2.16) transfers to a sphere in R^m. The positive-definite matrix V from (2.16) can be decomposed as

$$(3.3) \qquad V = C'C \, ,$$

where C is an upper triangular matrix with positive elements on the main diagonal

218

(so called Cholesky decomposition). If we define the transformation of the space R^m as

(3.4)
$$x \to x^* = Cx, \quad x \in R^m$$

(the asterisk will always denote transformed value) then the elipsoid $P(\alpha)$ will be obviously transformed to the form

(3.5)
$$P^*(\alpha) = \{b^* \in R^m : (b^* - \hat{b}^*)' (b^* - \hat{b}^*) \leqq k(\alpha)\},$$

which is an m-dimensional sphere with the center $\hat{b}^*$ and the radius $\sqrt{(k(\alpha))}$. The transformed solvability region S has the form

(3.6)
$$S^* = \{b^* \in R^m : \bar{h}_i' b^* \geqq 0, \ i = 1, ..., N\},$$

where

(3.7)
$$\bar{h}_i = (C^{-1})' h_i, \quad i = 1, ..., N.$$

If the interior of the elipsoid $P(\alpha)$ (denoted as int $P(\alpha)$) contains the zero vector 0 (or equivalently if the zero vector 0 lies in int $P^*(\alpha)$) then the problem discussed in this section has the following simple solution.

Lemma 1. Let $0 \in$ int $P(\alpha)$. Then the inclusion $P(\alpha) \subset S$ is true if and only if $S = R^m$.

Proof. The lemma is obvious since S is a cone with the vertex in 0 and $P(\alpha)$ is an elipsoid. □

General solution of the considered problem is given in the following theorem.

Theorem 1. The inclusion $P(\alpha) \subset S$ is true if and only if it holds

(3.8)
$$\bar{h}_i' \left(\hat{b}^* - \frac{\sqrt{(k(\alpha))}}{\|\bar{h}_i\|} \bar{h}_i \right) \geqq 0, \quad i = 1, ..., N,$$

where $\| \ \|$ is the usual Eucleidian norm in R^m.

Proof. $P(\alpha) \subset S$ is equivalent to $P^*(\alpha) \subset S^*$ and this last inclusion holds iff the sphere $P^*(\alpha)$ with the center $\hat{b}^*$ and the radius $\sqrt{(k(\alpha))}$ lies in all half-spaces $\{b^* \in R^m : \bar{h}_i' b^* \geqq 0\}$, $i = 1, ..., N$. This last condition is obviously equivalent to (3.8). □

Remark 3. The inequalities (3.8) can be written in the equivalent form

(3.9)
$$h_i' \hat{b} - \sqrt{(k(\alpha))} \|\bar{h}_i\| \geqq 0, \quad i = 1, ..., N$$

or

(3.10)
$$h_i' \hat{b} - \sqrt{(k(\alpha) \, h_i' V^{-1} h_i)} \geqq 0, \quad i = 1, ..., N.$$

4. OPTIMAL DECISIONS AND PREDICTION REGIONS

In this section some conclusions concerning the optimal decision in (1.1) with the random right-hand side will be shown which one can obtain from the prediction region for b using methods of parametric programming.

Let us denote

$$\varphi(b) = \min \{c'x: Ax = b, \ x \geqq 0\} \tag{4.1}$$

for $b \in S$. The function φ is convex, continuous and piecewise linear on S. More explicitly, there exist vectors $g_1, \dots, g_r \in R^m$ such that

$$\varphi(b) = \max \{g_j'b: j = 1, \dots, r\} . \tag{4.2}$$

According to the basis decomposition theorem (see [16]) the definition region S of the function φ can be decomposed to a finite number of convex polyhedral cones S_j with the vertices in the origin 0 such that the interiors of S_j are mutually disjunct and $\varphi(b) = g_j'b$ on S_j for $j = 1, \dots, r$. These cones correspond to particular bases B_j in A such that $\tilde{c}_j' B_j^{-1} A \leqq c'$ ($\tilde{c}_j$ is the subvector of c corresponding to B_j) and have the form

$$S_j = \{b \in R^m: B_j^{-1} b \geqq 0\} .$$

The cones S_j and the function φ can be found explicitly by means of the algorithms mentioned in [8] or [14].

a) First one can construct $(1 - \alpha)$ prediction interval for the optimal decision:

Theorem 2. Let $P(\alpha) \subset S$. Then it holds

$$\max \{\varphi(b): b \in P(\alpha)\} = \max_{j=1,\dots,r} \left\{ \bar{g}_j' \left(\hat{b}^* + \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|} \bar{g}_j \right) \right\}, \tag{4.3}$$

and

$$\min \{\varphi(b): b \in P(\alpha)\} \geqq \max_{j=1,\dots,r} \left\{ \bar{g}_j' \left(\hat{b}^* - \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|} \bar{g}_j \right) \right\}, \tag{4.4}$$

where

$$\bar{g}_j = (C^{-1})' g_j, \quad j = 1, \dots, r \tag{4.5}$$

and for $g_j = \bar{g}_j = 0$ we put

$$\bar{g}_j' \left(\hat{b}^* \pm \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|} \bar{g}_j \right) = 0 .$$

Proof. We can write

$$\max \{\varphi(b): b \in P(\alpha)\} = \max_{b \in P(\alpha)} \left\{ \max_{j=1 \dots r} \{g_j'b\} \right\} =$$

$$= \max_{b^* \in P^*(\alpha)} \left\{ \max_{j=1 \dots r} \{\bar{g}_j'b^*\} \right\} = \max_{j=1 \dots r} \left\{ \max_{b^* \in P^*(\alpha)} \{\bar{g}_j'b^*\} \right\} .$$

Now the relation (4.3) is proved since it obviously holds

$$\max_{b^* \in P^*(\alpha)} \{\bar{g}_j'b^*\} = \bar{g}_j' \left(\hat{b}^* + \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|} \bar{g}_j \right) \tag{4.6}$$

(we maximize the linear function $\bar{g}_j'b^*$ over the sphere with the center $\hat{b}^*$ and the radius $\sqrt{(k(\alpha))}$; the maximal value is achieved in the point where the vector directed from $\hat{b}^*$ as the gradient $\bar{g}_j$ of the function $\bar{g}_j'b^*$ crosses the surface of the sphere, i.e. in the point $\hat{b}^* + (\sqrt{(k(\alpha))}/\|\bar{g}_j\|) \bar{g}_j$).

As the relation (4.4) is concerned we have

$$\min\{\varphi(b): b \in P(\alpha)\} = \min_{b^* \in P^*(\alpha)}\{\max_{j=1,\dots,r}\{\bar{g}_j' b^*\}\} \geqq \max_{j=1,\dots,r}\{\min_{b^* \in P^*(\alpha)}\bar{g}_j' b^*\}.$$

The last inequality holds since it is

(4.7)
$$\max_{j=1,\dots,r}\{\bar{g}_j' b^*\} \geqq \max_{j=1,\dots,r}\{\min_{b^* \in P^*(\alpha)}\{\bar{g}_j' b^*\}\}$$

for each $b^* \in P^*(\alpha)$ so that we can replace the left-hand side of (4.7) by its minimal value over $b^* \in P^*(\alpha)$. The proof is finished because one can derive in the same way as (4.6) that

(4.8)
$$\min_{b^* \in P^*(\alpha)}\{\bar{g}_j' b^*\} = \bar{g}_j'\left(\hat{b}^* - \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|}\,\bar{g}_j\right). \qquad \square$$

Corollary. Let $P(\alpha) \subset S$. Then the optimal value of the objective function lies in the following interval

(4.9)
$$\left[\max_{j=1,\dots,r}\left\{\bar{g}_j'\left(\hat{b}^* - \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|}\,\bar{g}_j\right)\right\},\ \max_{j=1,\dots,r}\left\{\bar{g}_j'\left(\hat{b}^* + \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_j\|}\,\bar{g}_j\right)\right\}\right]$$

with probability at least $1 - \alpha$.

Remark 4. The interval (4.9) can be written again in the equivalent form

(4.10)
$$\left[\max_{j=1,\dots,r}\{g_j'\hat{b} - \sqrt{(k(\alpha)\,g_j'V^{-1}g_j)}\},\ \max_{j=1,\dots,r}\{g_j'\hat{b} + \sqrt{(k(\alpha)\,g_j'V^{-1}g_j)}\}\right].$$

Remark 5. Since the function $\varphi(b)$ attains the value $+\infty$ outside the set S (see e.g. [18]) one can omit the assumption $P(\alpha) \subset S$ in the previous corollary and formulate it in such a way that the optimal objective value lies in the interval (4.9) or is equal to $+\infty$ with the probability at least $1 - \alpha$.

Remark 6. This work is not the first one dealing with probability regions for optimal objective values in linear programs with random coefficients. E.g., results have been obtained by means of projection of rectangulars in which the values (A, B, c) lie with a given probability (see [12, Section 13.1] or [15]). These rectangulars are defined by means of the mean values and standard deviations of the random components of (A, b, c) and do not make use of the correlation structure (relations among particular random components) how it is the case when projecting elipsoids.

b) Further one can construct $(1 - \alpha)$ prediction region for the optimal decision:

According to (2.16) and the discussion from the beginning of this section it is obvious that $(1 - \alpha)$ prediction region for the optimal decision can be taken in the form

(4.11)
$$\bigcup_{j=1}^{r}\{x \in R^n\colon (\tilde{x}_j - B_j^{-1}\hat{b})'\, B_j'VB_j(\tilde{x}_j - B_j^{-1}\hat{b}) \leqq k(\alpha),\ \tilde{x}_j \geqq 0,\ \tilde{x}_j^- = 0\},$$

where $\tilde{x}_j$ denotes such subvector of the vector $x \in R^n$ which corresponds to the

basis B_j and $\tilde{x}_j^-$ contains the remaining components of x (it follows from (2.16) applied for particular optimal bases).

c) The approach based on $(1 - \alpha)$ prediction regions offers to define the following concept of $(1 - \alpha)$ decisions:

The vector $w \geqq 0$ in R^n is called $(1 - \alpha)$ decision if it fulfills the following two conditions

  (i) $Aw \in P(\alpha)$;

  (ii) **Prob** $\{ \min_{Ax=b, x \geqq 0} c'x < c'w \} \leqq \alpha.$

The probability condition (ii) means that a decision better than w may exist in our program with probability not higher than α.

The following obvious assertion suggests how to find $(1 - \alpha)$ decision if we know $P(\alpha)$.

Theorem 3. Let

(4.12) $$w = \operatorname{argmin} \{ c'x : Ax \in P(\alpha), \; x \geqq 0 \} .$$

Then w is $(1 - \alpha)$ decision.

5. APPLICATION OF BUNCHING METHOD

The practical applicability of the parametric programming methods from Sections 3 and 4 is limited if the dimensions of the problem are large and therefore more effective procedures are desirable which can be of an approximate character. One of such potential procedures based on so called bunching method is sketched in this section.

The method of bunching [19] and especially its trickling down modification in combination with Schur-complement bases updates [20] is the efficient tool for solving linear programs with variable right-hand sides (a bunch is such subset of a given set of right-hand side vectors which corresponds to the same optimal basis of the program).

In our case the bunching method will enable to solve in an efficient way a lot of problems of the type

(5.1) $$\{ \min c'x : Ax = z^k, \; x \geqq 0 \} , \quad k = 1, ..., K$$

(the points z^k are chosen from the elipsoid $P(\alpha)$) without performing explicitly the decomposition of the solvability region S to the cones S_j.

As the choice of the points z^k is concerned one can use various strategies. E. g., it is possible to choose the points z^k randomly from the surface of the elipsoid $P(\alpha)$. If we transform the coordinates according to (3.4) then we can generate these points

222

uniformly from the surface of the sphere (3.5) taking

$$(5.2) \qquad \begin{aligned} z_1^* &= \hat{b}_1^* + \sqrt{(k(\alpha))} \cos \vartheta_1 \cos \vartheta_2 \cos \vartheta_3 \dots \cos \vartheta_{m-1} \\ z_2^* &= \hat{b}_2^* + \sqrt{(k(\alpha))} \sin \vartheta_1 \cos \vartheta_2 \cos \vartheta_3 \dots \cos \vartheta_{m-1} \\ z_3^* &= \hat{b}_3^* + \sqrt{(k(\alpha))} \sin \vartheta_2 \cos \vartheta_3 \cos \vartheta_4 \dots \cos \vartheta_{m-1} \\ z_4^* &= \hat{b}_4^* + \sqrt{(k'(\alpha))} \sin \vartheta_3 \cos \vartheta_4 \cos \vartheta_5 \dots \cos \vartheta_{m-1} \\ &\;\;\vdots \\ z_{m-1}^* &= \hat{b}_{m-1}^* + \sqrt{(k(\alpha))} \sin \vartheta_{m-2} \cos \vartheta_{m-1} \\ z_m^* &= \hat{b}_m^* + \sqrt{(k(\alpha))} \sin \vartheta_{m-1}\,, \end{aligned}$$

where $0 \leqq \vartheta_1 \leqq 2\pi$, $-\pi/2 \leqq \vartheta_2 \leqq \pi/2, \dots,\ -\pi/2 \leqq \vartheta_{m-1} \leqq \pi/2$ are independent random variables with uniform distributions on their ranges.

The trickling down procedure can be started in the point $\hat{b}^*$ (the center of the sphere $P^*(\alpha)$). Let $B_{(1)}^*$ be the optimal basis for this point $\hat{b}^*$ (it holds obviously $B_{(1)}^* = C B_{(1)}$, where $B_{(1)}$ is the optimal basis for the point $\hat{b}$ before the transformation (3.4) since the problem (1.1) can be written equivalently as $\{\min c'x : CAx = Cb,\ x \geqq 0\}$). Let $z^{1*} = (z_1^{1*}, \dots, z_m^{1*})'$ be the first point generated according to (5.2). By using trickling down procedure (i.e. the proper sequence of dual simplex steps exploiting Schur-complement updates) one will find the corresponding sequence $B_{(1)}^*, \dots, B_{(s)}^*$ of the bases which is ended by the basis $B_{(s)}^*$ optimal for the point z^{1*}. Let us calculate the values

$$(5.3) \qquad l^1 = \bar{g}_{(s)}' \left(\hat{b}^* - \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_{(s)}\|}\, \bar{g}_{(s)} \right), \quad u^1 = \bar{g}_{(s)}' \left(\hat{b}^* + \frac{\sqrt{(k(\alpha))}}{\|\bar{g}_{(s)}\|}\, \bar{g}_{(s)} \right),$$

where

$$(5.4) \qquad \bar{g}_{(s)} = \left(B_{(s)}^{*-1} \right)' \tilde{c}_{(s)}$$

($\tilde{c}_{(s)}$ is the subvector of c corresponding to the basis $B_{(s)}^*$). The same procedure will be performed with the second generated point z^{2*} producing the values l^2 and u^2, etc. If proceeding in this way we obtain a tree rooted at the basis $B_{(1)}^*$ (see [20]) the paths of which are ended by the couples (l^1, u^1), (l^2, u^2), ..., (l^K, u^K). The $(1 - \alpha)$ prediction interval for the optimal objective value can be then approximated by the interval

$$(5.5) \qquad \left[\max_{k=1,\dots,K} l^k,\ \max_{k=1,\dots,K} u^k \right].$$

If it is $0 \in P(\alpha)$ then zero must be supplied to the numbers l^k for the calculation of the lower bound of the interval (5.5). The stopping rule by means of which number K is found can be prescribed in such a way that the last L couples (l^{K-L+1}, u^{K-L+1}), ..., (l^K, u^K) will satisfy

$$(5.6) \qquad \begin{aligned} \max_{k=1,\dots,K} l^k - \max_{k=1,\dots,L-K} l^k &< \varepsilon\,, \\ \max_{k=1,\dots,K} u^k - \max_{k=1,\dots,L-K} u^k &< \varepsilon\,, \end{aligned}$$

where an integer L and a sufficiently small $\varepsilon > 0$ are chosen apriori.

More complicated strategies can be suggested but the previous one seems to be suitable in spite of its simplicity. The inaccuracies which can originate when using the generating formulas (5.2) do not reduce the efficiency of the method since the points z^{1*}, z^{2*}, ... are used only to determine the corresponding optimal bases and these bases do not usually vary in the neighborhoods of particular right-hand side vectors. Moreover, when the components of the vector $\hat{b}$ are large (as it is frequent in practice) then usually only small number of the cones S_j from the decomposition of S have nonempty intersections with the elipsoid $P(\alpha)$ so that the mentioned tree from the trickling down procedure has small number of paths which reduces the computing effort.

Remark 5. We have described the construction of the prediction interval (5.5) for the optimal objective value. The method from this section also allows to construct approximate $(1 - \alpha)$ decisions defined in Section 4 and is applicable in many other situations. E.g. the prediction region $P(\alpha)$ may be rectangular of the form

$$P(\alpha) = \{b \in R^m : \alpha_i \leqq b_i \leqq \beta_i, \ i = 1, ..., m\}$$

(see Remark 2) or some coefficients of A and c may be also random (in this case we must have at our disposal a joint prediction region for all random coefficients).

6. NUMERICAL EXAMPLE

The authors of [14] investigated the following problem

(6.1)
$$\min \{x_2 + x_3 + 3x_4\}$$

$$\begin{aligned}
\text{s.t.} \quad x_1 - 2x_2 + x_3 - x_4 + x_5 &= b_1 \\
2x_1 + 3x_2 - x_3 + 2x_4 + x_6 &= b_2 \\
-x_1 + 2x_2 + 3x_3 - 3x_4 + x_7 &= b_3 \\
x_1, ..., x_7 &\geqq 0
\end{aligned}$$

In this case it is

(6.2)
$$S = \{b \in R^3 : 3b_1 + 2b_2 \geqq 0, \ 3b_2 + b_3 \geqq 0, \ b_1 + b_2 \geqq 0,$$
$$13b_1 + 8b_2 + b_3 \geqq 0, \ 3b_2 + 2b_3 \geqq 0\}$$

and

(6.3)
$$\varphi(b) = \max \{0, -\tfrac{1}{2}b_1, -b_2, -\tfrac{3}{4}b_1 - \tfrac{1}{4}b_3, -\tfrac{9}{8}b_1 - \tfrac{5}{8}b_3, -\tfrac{3}{4}b_2 - \tfrac{3}{2}b_3\}$$

for $b \in S$. Table 1 contains the description of all cones S_j from the decomposition of S including the corresponding forms of φ and the optimal bases B_j.

Let the 95 per cent prediction region (2.16) have the following form

(6.4)
$$\begin{pmatrix} b_1 - 1\cdot2 \\ b_2 - 8\cdot4 \\ b_3 + 7\cdot6 \end{pmatrix}' \begin{pmatrix} 6\cdot25 & -3 & 0\cdot75 \\ -3 & 19\cdot93 & -9\cdot39 \\ 0\cdot75 & -9\cdot39 & 17\cdot46 \end{pmatrix} \begin{pmatrix} b_1 - 1\cdot2 \\ b_2 - 8\cdot4 \\ b_3 + 7\cdot6 \end{pmatrix} \leqq 6\cdot76,$$

Table 1. The analysis of the linear parametric problem (6.1).

i	S_i	$\varphi(b)$	B_j (numbers of columns of A)
1	$b_1 \geqq 0,\ b_2 \geqq 0,\ b_3 \geqq 0$	0	(5, 6, 7)
2	$-b_3 \geqq 0,\ b_1 + b_3 \geqq 0,$ $b_2 + 2b_3 \geqq 0$	0	(1, 5, 6)
3	$-b_1 \geqq 0,\ 3b_1 + 2b_2 \geqq 0,$ $b_1 + b_3 \geqq 0$	$-\frac{1}{2}b_1$	(2, 6, 7)
4	$-b_2 \geqq 0,\ b_1 + b_2 \geqq 0,$ $3b_2 + b_3 \geqq 0$	$-b_2$	(3, 5, 7)
5	$-b_1 - b_3 \geqq 0,\ 3b_1 - b_3 \geqq 0,$ $-b_1 + b_2 + b_3 \geqq 0$	$-\frac{3}{4}b_1 - \frac{3}{4}b_3$	(1, 4, 6)
6	$-b_1 - b_3 \geqq 0,\ -3b_1 + b_3 \geqq 0,$ $13b_1 + 8b_2 + b_3 \geqq 0$	$-\frac{9}{8}b_1 - \frac{5}{8}b_3$	(2, 4, 6)
7	$-b_2 - 2b_3 \geqq 0,\ 3b_2 + 2b_3 \geqq 0,$ $b_1 - b_2 - b_3 \geqq 0$	$-\frac{3}{4}b_2 - \frac{3}{2}b_3$	(1, 4, 5)

i.e.

$$(6.5) \qquad \hat{b} = \begin{pmatrix} 1\cdot2 \\ 8\cdot4 \\ -7\cdot6 \end{pmatrix}, \quad V = \begin{pmatrix} 6\cdot25 & -3 & 0\cdot75 \\ -3 & 19\cdot93 & -9\cdot39 \\ 0\cdot75 & -9\cdot39 & 17\cdot46 \end{pmatrix}, \quad C = \begin{pmatrix} 2\cdot5 & -1\cdot2 & 0\cdot3 \\ 0 & 4\cdot3 & -2\cdot1 \\ 0 & 0 & 3\cdot6 \end{pmatrix}$$

According to (6.2) the vectors h_i from (3.2) are

$$(6.6) \qquad h_1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \quad h_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \quad h_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad h_4 = \begin{pmatrix} 13 \\ 8 \\ 1 \end{pmatrix}, \quad h_5 = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$$

and according to Table 1 the vectors g_j from (4.2) are

$$(6.7) \qquad g_1 = g_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad g_3 = \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix}, \quad g_4 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \quad g_5 = \begin{pmatrix} -\frac{3}{4} \\ 0 \\ -\frac{3}{4} \end{pmatrix},$$

$$g_6 = \begin{pmatrix} -\frac{9}{8} \\ 0 \\ -\frac{5}{8} \end{pmatrix}, \quad g_7 = \begin{pmatrix} 0 \\ -\frac{3}{4} \\ -\frac{3}{2} \end{pmatrix}.$$

The left-hand sides of the inequalities (3.8) (or equivalently of (3.9) or (3.10)) are

$$(6.8) \qquad\qquad 14\cdot48,\ 13\cdot71,\ 7\cdot40,\ 49.70,\ 5\cdot27\ .$$

Since each of these values is non-negative the elipsoid (6.4) is the subset of the solvability region (6.2). The corresponding 95 per cent prediction interval for the optimal value of the objective function is according to (4.9) or (4.10)

$$(6.9) \qquad\qquad\qquad [3\cdot256,\ 7\cdot274]\ .$$

If using the bunching method 300 points generated according to (5.2) gave the following 95 per cent prediction interval for the optimal objective value (see (5.5))

$$(6.10) \qquad\qquad\qquad [3\cdot338,\ 7\cdot271]$$

225

which is comparable with the exact interval (6.9). Moreover, we have obtained 95 per cent decision of the form

$$(6.11) \qquad x_1 = 3{\cdot}403\,, \quad x_4 = 1{\cdot}113\,, \quad x_6 = 0{\cdot}101\,, \quad x_2 = x_3 = x_5 = x_7 = 0\,.$$

REFERENCES

[1] G. E. P. Box and G. M. Jenkins: Time Series Analysis. Forecasting and Control. Holden Day, San Francisco 1970.
[2] T. Cipra: Time Series Analysis with Applications in Economy (in Czech). SNTL/ALFA, Prague 1986.
[3] T. Cipra: Confidence regions for linear programs with random coefficients. Working Paper WP-86-20, International Institute for Applied Systems Analysis, Laxenburg 1986.
[4] T. Cipra: Time series in linear programs with random right-hand sides. Working Paper WP-86-21, International Institute for Applied Systems Analysis, Laxenburg 1986.
[5] T. Cipra: Time series in linear programs. International Conference on Stochastic Programming, Prague 1986 (unpublished).
[6] P. J. Dhrymes: Econometrics. Statistical Foundations and Applications. Springer, New York 1974.
[7] J. Dupačová and R. Wets: Asymptotic behavior of statistical estimators and optimal solutions for stochastic optimization problems. Working Paper WP-86-41, International Institute for Applied Systems Analysis, Laxenburg 1986.
[8] T. Gal and J. Nedoma: Multiparametric linear programming. Management Sci. *18* (1972), 406—422.
[9] C. W. J. Granger and P. Newbold: Forecasting Economic Time Series. Academic Press, New York 1977.
[10] S. H. Hymans: Simultaneous confidence intervals in econometric forecasting. Econometrica *36* (1968), 18—30.
[11] J. Johnston: Econometric Methods. McGraw-Hill, New York 1972.
[12] D. B. Judin: Mathematical Methods of Control under Incomplete Information — Problems and Methods of Stochastic Programming (in Russian). Sovietskoje Radio, Moskva 1974.
[13] D. C. Montgomery and L. A. Johnson: Forecasting and Time Series Analysis. McGraw-Hill, New York 1976.
[14] F. Nožička, J. Guddat, H. Hollatz and B. Bank: Theorie der linearen parametrischen Optimierung. Akademie-Verlag, Berlin 1974.
[15] R. Theodorescu: Random programs. Math. Operationsforsch. Statist. *3* (1972), 19—47.
[16] D. Walkup and R. Wets: Lifting projections of convex polyhedra. Pacific J. Math. *28* (1969), 465—475.
[17] R. Wets: Programming under uncertainty: the equivalent convex program. SIAM J. Appl. Math. *14* (1966), 89—105.
[18] R. Wets: Stochastic programs with fixed recourse: the equivalent deterministic program. SIAM Review *16* (1974), 309—339.
[19] R. Wets: Stochastic programming: solution techniques and approximation schemes. In: Mathematical Programming: The State of Art (A. Bachem, M. Grotsched, B. Korte, eds.). Springer, Berlin 1983, 566—603.
[20] R. Wets: Large scale linear programming techniques in stochastic programming. Working Paper WP-84-90, International Institute for Applied Systems Analysis, Laxenburg 1984.

RNDr. Tomáš Cipra, CSc., matematicko-fyzikální fakulta UK (Faculty of Mathematics and Physics - Charles University), Sokolovská 83, 186 00 Praha 8. Czechoslovakia.