2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Recursive Estimation Algorithm for l_1-norm Approximation in Dynamic Systems with Nonoptimality Levels*

Pavel Akimov and Alexander Matasov[1]

Abstract— **State estimation problem for a linear discrete dynamic system is considered. Some components of the state vector can abruptly change under the influence of rare uncontrolled input pulses in the right-hand side of equations. In this case, l_1-norm approximation (least absolute deviations method) gives better results than the standard l_2-norm approximation (least squares method). A recursive estimation algorithm for finding l_1-norm approximation in case of large amount of measurements is presented. To make numerical procedure more reliable, a nonoptimality level for current iteration is constructed. An example from inertial navigation verifies the effectiveness of proposed approach.**

I. INTRODUCTION

The approximation using l_1-norm (least absolute deviations method) is very effective for processing the measurement data with outliers [1]-[3] and also in case when a part of state vector components can have stepwise jumps under the influence of rare uncontrolled input pulses [4], [5]. The main aim of our investigation is to elaborate a recursive numerical algorithm for solving the problem of l_1-norm approximation for dynamic systems with large measurement arrays.

In the previous articles [4] and [5], the dynamic estimation problems were reduced to static problems with an unknown parameter vector of high dimension. In the present paper, we investigate a method that takes into account the dynamic structure of estimation problem. This allows us to avoid the crucial dependence on the amount of measurement data. The method is based on successive solution of quadratic problems with special selection of weight coefficients. Applied to static problems, a similar approach (the algorithm of variational-weighted quadratic approximations) is implemented in practice; in [3] it was also called Weiszfeld's method. For dynamic systems, this method can be improved through the use of the recurrent relations between the estimates of state vector at different time instants. This allows us to substantially increase the reliability of numerical procedures. An important aspect related to iterative solving algorithms is the necessity of evaluating the performance of the approximate solution obtained at current iteration. Below we construct a nonoptimality level of current iteration that takes into account the dynamic character of the problem. Thus the present investigation is a significant development of our previous results.

*This work was supported by the RFBR Grant No.11-08-00004-a

[1]P. Akimov and A. Matasov are with the Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Leninskie Gory, Moscow GSP-1 119991, Russia akmpavel at rambler.ru, alexander.matasov at gmail.com

Potentialities of the proposed approach are illustrated by numerical examples from the theory of bench tests for strapdown inertial navigation systems.

II. APPROXIMATION PROBLEMS

A. Problems of l_1-norm and l_2-norm approximation

Consider a dynamic system that is described by equation

$$X(k+1) = F\,X(k) + G\,q(k) + g(k), \qquad k = 0, \ldots, K-1,$$

where $X(k) \in \mathbf{R}^n$ is an unknown state vector at instant k, $q(k) \in \mathbf{R}^l$ is an unknown input disturbance vector, $F \in \mathbf{R}^{n\times n}$, $G \in \mathbf{R}^{n\times l}$ are specified matrices, and $g(k) \in \mathbf{R}^n$ is a known vector of system nonhomogeneity. It is assumed for simplicity that the system matrices are constant. All statements and conclusions that are discussed below are valid for time-varying matrices as well.

The problem of estimating the state vector $X(k)$ will be solved in deterministic setting: we make no assumptions concerning statistical properties of $q(k)$. Suppose only that the scales (typical magnitudes) of its elements are specified:

$$q_\beta(k) \sim Q_\beta, \qquad \beta = 1, \ldots, l, \quad k = 0, \ldots, K-1.$$

Suppose we also have a priori estimate $\bar{X}(0)$ for initial system state $X(0)$:

$$\bar{X}(0) = X(0) + \bar{r}(0),$$

where $\bar{r}(0) \in \mathbf{R}^n$ is an error vector. The scales of $\bar{r}_\alpha(0)$ are specified by values Π_α:

$$\bar{r}_\alpha(0) \sim \Pi_\alpha, \quad \alpha = 1, \ldots, n.$$

Let the measurements be performed at each time instant k:

$$z(k) = HX(k) + r(k), \qquad k = 0, \ldots, K.$$

Here $z(k) \in \mathbf{R}^m$ is the measurement vector, $H \in \mathbf{R}^{m\times n}$ is a specified matrix, $r(k) \in \mathbf{R}^m$ is a measurement noise.

As for system disturbances, we assume (with no statistical hypotheses) that the typical magnitudes of measurement noises are known:

$$r_\gamma(k) \sim R_\gamma, \qquad \gamma = 1, \ldots, m, \quad k = 0, \ldots, K.$$

At some rare instants the anomalous values of disturbances $q(k)$ can lead to noticeable jumps (of stepwise type) in a part of state vector components.

Introduce auxiliary diagonal matrices

$$\begin{aligned} \Pi^{-1} &= \operatorname{diag}\left(\Pi_1^{-1}, \ldots, \Pi_n^{-1}\right), \\ Q^{-1} &= \operatorname{diag}\left(Q_1^{-1}, \ldots, Q_l^{-1}\right), \\ R^{-1} &= \operatorname{diag}\left(R_1^{-1}, \ldots, R_m^{-1}\right). \end{aligned}$$

In order to obtain the estimates for state vectors we consider the following variational problem for l_1-norm approximation:

$$I_0 = \min_{X,\,q} \left\|\Pi^{-1}(\bar{X}(0) - X(0))\right\|_1 + \sum_{k=0}^{K-1} \left\|Q^{-1}q(k)\right\|_1 \quad (1)$$
$$+ \sum_{k=0}^{K} \left\|R^{-1}(z(k) - HX(k))\right\|_1$$

subject to

$$X(k+1) - F\,X(k) - G\,q(k) - g(k) = 0 \qquad (2)$$

for $k = 0, \ldots, K-1$. As usual, $\|\cdot\|_1$ means l_1-norm: $\|y\|_1 = \sum_s |y_s|$. The notation $\{X,\, q\}$ is an abbreviation for $\{X(0), \ldots, X(K),\, q(0), \ldots, q(K-1)\}$.

Therefore, we must minimize the sum of absolute values of data residuals under linear constraints of equality type. The presence of these constraints generates a difference between variational problem (1), (2) and the classical least absolute deviations method [1],[3].

More common than (1), (2) smoothing problem is the quadratic problem of the following form (l_2-norm approximation) [6]:

$$J_0 = \min_{X,\,q} \left(\left\|\Pi^{-1}(\bar{X}(0) - X(0))\right\|_2^2 \right. \qquad (3)$$
$$\left. + \sum_{k=0}^{K-1} \left\|Q^{-1}q(k)\right\|_2^2 + \sum_{k=0}^{K} \left\|R^{-1}(z(k) - HX(k))\right\|_2 \right)^{\frac{1}{2}}$$

under the same constraints (2).
Here $\|\cdot\|_2$ means l_2-norm: $\|y\|_2 = \left(\sum_s y_s^2\right)^{\frac{1}{2}}$.

There exist effective numerical methods for solving problem (3), (2). A basic analytical result (Bryson-Frazier formulas) is presented in the following statement [6].

THEOREM 1. *A solution* $\{X^*(0), \ldots, X^*(K),\, q^*(0), \ldots, q^*(K-1)\}$ *for problem (3), (2) is defined by the formulas:*

$$X^*(j) \;=\; X^-(j) + P^-(j)\lambda^J(j), \qquad j = 0, \ldots, K,$$
$$q^*(j) \;=\; Q^2 G^T \lambda^J(j+1), \qquad j = 0, \ldots, K-1,$$

where $X^-(j)$ *and* $P^-(j)$ *are the state vector estimate and the covariance matrix for the Kalman filter, respectively:*

$$X^-(j+1) \;=\; FX^-(j) + K_p(j)\,(z(j) - HX^-(j)) + g(j),$$
$$P^-(j+1) \;=\; F\,P^-(j)\,F^T + G\,Q^2\,G^T$$
$$- K_p(j)\,(R^2 + HP^-(j)H^T)\,K_p^T(j),$$
$$K_p(j) \;=\; F\,P^-(j)\,H^T\,(R^2 + HP^-(j)H^T)^{-1},$$
$$j = 0, \ldots,\ K\text{–}1, \qquad X^-(0) = \bar{X}(0), \qquad P^-(0) = \Pi^2.$$

The vector $\lambda^J(j)$ *is defined by the backwards-time recursions:*

$$\lambda^J(j) \;=\; (F - K_p(j)H)^T\,\lambda^J(j+1)$$
$$+ H^T\,(R^2 + HP^-(j)H^T)^{-1}\,(z(j) - HX^-(j)),$$
$$\lambda^J(K+1) = 0, \qquad j = 0, \ldots, K.$$

REMARK 1. Strictly speaking, in the book [6] Theorem 1 is proved for the case of $\bar{X}(0) = 0$ and $g(k) = 0$. However it is easy to verify that Theorem 1 directly follows from the results of [6]. To do it, one should proceed to the consideration of the centered process $X(k) - \bar{X}(k)$, where vectors $\bar{X}(k)$ are known and are described by the recurrent formulas

$$\bar{X}(k+1) = F\,\bar{X}(k) + g(k), \qquad k = 0, \ldots, K-1,$$

with known initial value $\bar{X}(0)$.

The Bryson-Frazier formulas allow us to find the solution of quadratic smoothing problem by means of two recursive passes. Another important aspect of numerical implementation is the application of the square-root method [6].

The standard numerical algorithms for problem (1), (2) are based on linear programming (LP), interior point method [2]. However they operate with matrices of order $(Km + Kl) \times (Kn + Km + Kl)$ that requires substantial computational resources under large K (e.g., of the order of several thousand). In many applications, e.g., when processing navigation data, long-lasting series of measurements can be considered. In these cases, another methods are necessary that use dynamic nature of estimation problem and avoid the reduction to static problem with high dimension of vectors and matrices. Below, we examine one such approach.

B. Method of Weight and Time Recurrences

In this section, we modify the algorithm of variational-weighted quadratic approximations (Weiszfeld's algorithm) from static case [3], [5] to dynamic one. The main difference is how l_2-norm approximation problems are solved at each iteration of Weiszfeld's algorithm, and thus in what form the corresponding dual problems (see Section III below) are presented.

Let $\{X(k,0),\ q(j,0);\ k = 0, \ldots, K,\ j = 0, \ldots, K-1\}$ be a feasible element for problem (1), i.e., it satifies (2). In this notation, the first argument in parentheses is a time instant, and the second argument (in this case, 0) is a number of iteration. Consider a sequence of iterations; at each iteration s we seek a minimum of quadratic function (3) with the weight coefficients that depend on the approximate solution obtained at previous iteration:

$$J_{W0} = \min_{X,\,q} \left((\bar{X}(0) - X(0))^T\,\Pi_W^{-2}(s)\,(\bar{X}(0) - X(0)) \right.$$
$$+ \sum_{k=0}^{K-1} q^T(k)\,Q_W^{-2}(k,s)\,q(k) \qquad (4)$$
$$\left. + \sum_{k=0}^{K} (z(k) - HX(k))^T R_W^{-2}(k,s)(z(k) - HX(k)) \right)^{\frac{1}{2}}$$

subject to (2).

Here $\Pi_W^{-2}(s), Q_W^{-2}(j,s), R_W^{-2}(k,s)$ are diagonal weight matrices with the entries that depend on the instant of time

and the solution at previous iteration $s = 0, 1, \ldots$:

$$\Pi^{-2}_{W\alpha}(s)=\begin{cases}\dfrac{\Pi^{-1}_{\alpha}}{|\bar{X}_{\alpha}(0)-X_{\alpha}(0,s)|} & \text{if } |\bar{X}_{\alpha}(0)-X_{\alpha}(0,s)|>\theta\Pi_{\alpha},\\ \Pi^{-2}_{\alpha}/\,\theta & \text{if } |\bar{X}_{\alpha}(0)-X_{\alpha}(0,s)|\le\theta\Pi_{\alpha},\end{cases}$$

$$Q^{-2}_{W\beta}(j,s)=\begin{cases}Q^{-1}_{\beta}/\,|q_{\beta}(j,s)| & \text{if } |q_{\beta}(j,s)|>\theta Q_{\beta},\\ Q^{-2}_{\beta}/\,\theta & \text{if } |q_{\beta}(j,s)|\le\theta Q_{\beta},\end{cases}$$

$$R^{-2}_{W\gamma}(k,s)=\begin{cases}\dfrac{R^{-1}_{\gamma}}{|z_{\gamma}(k)-H_{\gamma}X(k,s)|} & \text{if } |z_{\gamma}(k)-H_{\gamma}X(k,s)|>\theta R_{\gamma},\\ R^{-2}_{\gamma}/\,\theta & \text{if } |z_{\gamma}(k)-H_{\gamma}X(k,s)|\le\theta R_{\gamma},\end{cases}$$

where $\{X(k,s), q(j,s)\}$ is a solution obtained at previous iteration with number s. The dimensionless factor θ characterizes the smallness of residuals. Thus, e.g., instead of the term $R^{-1}_{\gamma}|z_{\gamma}(k) - H_{\gamma}X(k)|$ in functional (1) there appears the summand

$$\frac{R^{-1}_{\gamma}(z_{\gamma}(k) - H_{\gamma}X(k))^2}{|z_{\gamma}(k) - H_{\gamma}X(k,s)|},$$

where we use the approximation $|(\cdot)| \approx (\cdot)^2 / |(\cdot)_{\text{approx}}|$. Besides, the regularization is performed: if at some iteration a residual becomes too small, i.e.,

$$|z_{\gamma}(k) - H_{\gamma}X(k,s)| \le \theta R_{\gamma},$$

then the corresponding coefficient is fixed. Similar procedure is carried out for other components.

A solution of quadratic problem (4), (2) for next iteration $s+1$ is found by means of recursive Bryson-Fraizer formulas (see Theorem 1) improved by the square-root method [6]. Certainly, Theorem 1 is valid for time-varying weight matrices, i.e., for $\Pi \to \Pi(s)$, $Q \to Q_W(j,s)$, and $R \to R_W(k,s)$.

This recursiveness allow us to deal with vectors and matrices of relatively small dimension (of order n and $n\times n$, respectively), that substantially saves computational resources and computer memory compared with the reduction to static case [4], where we have to operate with an unknown parameter vector of dimension $n + Kl$ and corresponding matrices of large order $(n + (K + 1)m + Kl) \times (n + Kl)$.

In what follows, we assume that the minimal value of quadratic functional at each iteration does not equal to zero. If this is not the case, and at some iteration the problem value equals zero, then every summand in this functional equals zero and all residuals at this iteration equal zero. Consequently $I_0 = 0$ and the solution for initial smoothing problem (1), (2) is obtained at the above mentioned iteration.

We consider a solution of quadratic problem (4), (2) obtained at each iteration s as an approximate solution to initial nonsmooth problem (1), (2). So, we have two embedded iterative processes: in the outer loop, the auxiliary quadratic smoothing problems (4), (2) are formed basing on the previous step of this loop; in the inner loop, a solution of each such problem is found by means of Theorem 1. In the first case, we talk about iterations parameterized by s, in the second case, we consider recursions in time.

C. Nonoptimality Levels

The described above algorithm is iterative. Moreover, there is no rigorous result on its convergence; only successful numerical experience verifies its good performance. Therefore it is very desirable to evaluate iteration accuracy.

Further, in the notation of weight coefficients and the solution of (4), (2) we will omit the number of current iteration s supposing

$$\Pi_W(s)=\Pi_W, \quad Q_W(j,s)=Q_W(j), \quad R_W(k,s)=R_W(k),$$
$$X(k,s+1) = X^*(k), \quad q(j,s+1) = q^*(j)$$

since all conclusions are valid for any iteration.

We characterize the accuracy of the iteration obtained at the s-th step by the nonoptimality level (cf. [7])

$$\Delta = I(X^*,\, q^*)/I_0, \tag{5}$$

where I_0 is the unknown minimal value of the nonsmooth functional from (1) and the quantity in the numerator

$$I(X^*,\, q^*)=I\,(X^*(0),\ldots,X^*(K),q^*(0),\ldots,q^*(K-1))$$

is the value of the functional from (1) at current iteration. Obviously, $\Delta \ge 1$; the closer to 1 is Δ, the more accurate is the approximate solution. Since I_0 is unknown, Δ is also unknown. However it is possible to construct an upper bound for Δ: $\Delta \le \Delta_0$, where Δ_0 can actually be computed (see Section III); Δ_0 can be regarded as a guaranteed nonoptimality level. If Δ_{end} is a given threshold, then the calculation process should be halted when $\Delta_0 \le \Delta_{\text{end}}$.

We now discuss the choice of Δ_{end}. Let, for each summand in functional (1), we can (approximately) specify a value ϵ that defines a tolerable, not significant deviation from the optimal value. For physically well-defined variational problems the scaling is carried out, and the typical magnitudes of all summands are not much different. Therefore, the value of ϵ can be chosen the same for all summands.

So, let $\{X^0(k),\, q^0(j)\}$ and $\{X^*(k),\, q^*(j)\}$, where $k = 0, \ldots, K,\ j = 0, \ldots, K-1$, be the unknown solution of (1), (2) and a successful approximation for this solution, respectively, i.e., the following inequalities hold:

$$|\Pi^{-1}_{\alpha}\,(\bar{X}_{\alpha}(0) - X^*_{\alpha}(0))| - |\Pi^{-1}_{\alpha}\,(\bar{X}_{\alpha}(0) - X^0_{\alpha}(0))| \le \epsilon,$$
$$|Q^{-1}_{\beta}\,q^*_{\beta}(j)| - |Q^{-1}_{\beta}\,q^0_{\beta}(j)| \le \epsilon, \tag{6}$$
$$|R^{-1}_{\gamma}\,(z_{\gamma}(k) - H_{\gamma}X^*(k))| - |R^{-1}_{\gamma}\,(z_{\gamma}(k) - H_{\gamma}X^0(k))| \le \epsilon.$$

Summing up both sides of (6) we get the inequality

$$I(X^*,\, q^*) - I_0 \le (n + (K + 1)m + Kl)\epsilon$$

whence

$$\Delta \le \left(1 - \frac{(n + (K + 1)m + Kl)\epsilon}{I(X^*,\, q^*)}\right)^{-1}.$$

Hence the stopping criterion has the form

$$\Delta_0 \le \Delta_{\text{end}}, \qquad \Delta_{\text{end}} = \left(1 - \frac{(n + (K + 1)m + Kl)\epsilon}{I(X^*,\, q^*)}\right)^{-1}.$$

In order to construct the desired guaranteed level Δ_0 we must consider dual problems for primal problems (1), (2) and (3), (2) (or, that is essentially the same, (4), (2)).

III. DUAL PROBLEMS FOR CONSTRUCTING THE NONOPTIMALITY LEVELS

A. Auxiliary Dual Problems

Using cumbersome but standard calculation [8], we can construct dual problems to the primal problems considered above.

THEOREM 2. *The dual problem to (1), (2) has the form:*

$$I^0=\sup_{\mu,\,\lambda}\left(\bar{X}^T(0)\mu_0+\sum_{k=0}^{K}z^T(k)\mu_z(k)+\sum_{k=0}^{K-1}g^T(k)\lambda(k+1)\right) \tag{7}$$

subject to (as usual, $\|y\|_\infty=\max_s|y_s|$)

$$\max\left(\|\Pi\mu_0\|_\infty,\|\{Q\mu_q(j)\}_{j=0}^{K-1}\|_\infty,\|\{R\mu_z(k)\}_{k=0}^{K}\|_\infty\right)\le 1, \tag{8}$$

$$\begin{aligned}\mu_0-F^T\lambda(1)+H^T\mu_z(0) &= 0,\\ \lambda(k)-F^T\lambda(k+1)+H^T\mu_z(k) &= 0,\\ -G^T\lambda(k)+\mu_q(k-1) &= 0,\\ k=1,\dots,K,\qquad \lambda(K+1) &= 0.\end{aligned} \tag{9}$$

Moreover, the duality relation holds: $I_0=I^0$.

THEOREM 3. *The dual problem to (4), (2) has the form:*

$$J_W^0=\sup_{\mu,\,\lambda}\left(\bar{X}^T(0)\mu_0+\sum_{k=0}^{K}z^T(k)\mu_z(k)+\sum_{k=0}^{K-1}g^T(k)\lambda(k+1)\right) \tag{10}$$

subject to

$$\left(\mu_0^T\Pi_W^2\mu_0+\sum_{k=0}^{K-1}\mu_q^TQ_W^2(k)\mu_q(k)+\sum_{k=0}^{K}\mu_z^T(k)R_W^2(k)\mu_z(k)\right)^{\frac{1}{2}}\le 1 \tag{11}$$

and constraints (9). Moreover, the duality relation holds: $J_{W0}=J_W^0$.

Applying traditional methods of the theory of convex variational problems [8], one can establish the relation between the solutions of primal problem (4), (2) and dual problem (10), (11), and (9).

THEOREM 4. *Let $\{X^*,q^*\}$ be a solution of problem (4), (2), and J_{W0} be the minimal value of functional (4). Then the solution $\{\mu^*,\lambda^*\}$ of dual problem (10), (11), (9) is defined by the following relations:*

$$\begin{aligned}\mu_0^* &= J_{W0}^{-1}\cdot\Pi_W^{-2}\,(\bar{X}(0)-X^*(0)),\\ \mu_z^*(k) &= J_{W0}^{-1}\cdot R_W^{-2}(k)\,(z(k)-HX^*(k)),\quad k=0,\dots,K,\\ \mu_q^*(j) &= -J_{W0}^{-1}\cdot Q_W^{-2}(j)\,q^*(j),\quad j=0,\dots K-1,\\ \lambda^*(k) &= -J_{W0}^{-1}\cdot\lambda^J(k),\qquad k=1,\dots,K,\end{aligned}$$

where the vectors $\lambda^J(k)$ are found by the backwards recursion:

$$\lambda^J(k)=F^T\lambda^J(k+1)+H^TR_W^{-2}(k)\,(z(k)-HX^*(k)),$$
$$k=0,\dots,K,\qquad \lambda^J(K+1)=0.$$

By means of Theorem 4 we can easily calculate the optimal dual vector via the solution of primal problem.

B. Construction of Guaranteed Nonoptimality Level

To obtain the upper bound for nonoptimality level (5) (i.e., a guaranteed nonoptimality level Δ_0) one should evaluate the unknown I_0 from above. Due to Theorem 2, the duality relation $I_0=I^0$ holds. Therefore, for any feasible dual vector (μ,λ) satisfying (8) and (9), we get

$$I_0\ge\left(\bar{X}^T(0)\mu_0+\sum_{k=0}^{K}z^T(k)\mu_z(k)+\sum_{k=0}^{K-1}g^T(k)\lambda(k+1)\right). \tag{12}$$

Note that primal problem (4), (2) approximates primal problem (1), (2). Similarly, dual problem (10), (11), (9) approximates dual problem (7), (8), (9). Hence the known (due to Theorem 4) solution (μ^*,λ^*) of (10), (11), (9) suggests a direction of search. Namely, let us search for a successful feasible dual vector on the line $\sigma(\mu^*,\lambda^*)$, where σ is a scalar. Then a desirable lower bound for I_0 takes the form

$$I_0\ge\sup_\sigma\sigma\left(\bar{X}^T(0)\mu_0^*+\sum_{k=0}^{K}z^T(k)\mu_z^*(k)+\sum_{k=0}^{K-1}g^T(k)\lambda^*(k+1)\right)$$

under the scalar constraint

$$|\sigma|\cdot\|\,(\,\Pi\mu_0^*,\,\{Q\mu_q^*(j)\}_{j=0}^{K-1},\,\{R\mu_z^*(k)\}_{k=0}^{K})\,\|_\infty\le 1.$$

Obviously, the latter extremal problem can be easily resolved. Full implementation of the described above idea leads to the main result of Section III, which is presented in the following assertion.

THEOREM 5. *Let $\{X^*,q^*\}$ be a solution of problem (4), (2) for l_2-norm approximation. Then $\Delta\le\Delta_0$, where*

$$\Delta_0=I(X^*,\,q^*)\cdot\theta_{\max}\,/\,J_{W0}^2.$$

Here the values $I(X^*,\,q^*)$, J_{W0}, and $\theta_{\max}$ are determined by the following relations:

$$I(X^*,\,q^*)=\left\|\Pi^{-1}(\bar{X}(0)-X^*(0))\right\|_1+\sum_{k=0}^{K-1}\left\|Q^{-1}q^*(k)\right\|_1+\sum_{k=0}^{K}\left\|R^{-1}(z(k)-HX^*(k))\right\|_1,$$

$$J_{W0}=\left(\left\|\Pi_W^{-1}(\bar{X}(0)-X^*(0))\right\|_2^2+\sum_{k=0}^{K-1}\left\|Q_W^{-1}(k)q^*(k)\right\|_2^2+\sum_{k=0}^{K}\left\|R_W^{-1}(k)(z(k)-HX^*(k))\right\|_2^2\right)^{\frac{1}{2}},$$

and

$$\theta_{\max}=\max\Big(\,\|\Pi\cdot\Pi_W^{-2}\left(\bar{X}(0)-X^*(0)\right)\|_\infty,\ \|\{Q\cdot Q_W^{-2}(j)q^*(j)\}_{j=0}^{K-1}\|_\infty,\ \|\{R\cdot R_W^{-2}(k)(z(k)-HX^*(k))\}_{k=0}^{K}\|_\infty\,\Big).$$

Theorem 5 allows us to constructively evaluate the nonoptimality of current solution at each iteration of estimation algorithm. Thereby we get a criterion for halting the iteration process. In [5] we obtained a similar result but for the static case without dynamic constraints (2). So, Theorem 5 extends this result to dynamic case.

IV. NUMERICAL IMPLEMENTATION

A. Problem Description

Engineers set us the following problem from inertial navigation. The biases of inertial sensors (accelerometers and gyroscopes), as a rule, vary insignificantly for long time periods. Therefore, before the admission of the strapdown inertial navigation system (SDINS) to the operation, the bench tests are performed to determine the biases in order to compensate their influence. However, sometimes, rare jumps (of step type) can occur in sensor biases. In other words, the average values of sensor noises can change noticeably. In this case, the calibration is not possible. So, it is necessary to determine the magnitudes and instants of these jumps. The approach based on l_1-norm approximation turns out to be very effective for solving this engineering problem, in contrast to the pure l_2-norm algorithm (least squares method) [4].

As a model, consider the error equations of SDINS under bench testing [4]. This model consists of the equations for dynamic system, which describe the variation of navigation errors, the equations for measurements during test, and the prior information, which, in the simplest case, is a knowledge of the scales for state vector entries at initial instant.

Note that the formal definition of jump exists for continuous-time signals. The initial differential equations for the errors of sensors have the form

$$\dot{\kappa}_i(t) = q_i(t), \qquad i = 1, \ldots, 4,$$

where κ_1, κ_2 are the errors of accelerometers, and κ_3, κ_4 are the errors of gyroscopes; q_i are disturbances. In these equations, the biases are modeled by (unknown) initial conditions, the high-frequency noises are modeled by the high-frequency components of $q_i(t)$, and the (possible) jumps are modeled by pulses in $q_i(t)$.

The state vector in the corresponding discrete-time model has the form $X = (\delta V_1, \delta V_2, \beta_1, \beta_2, \kappa_1, \ldots, \kappa_4)^T$. Here, δV_1, δV_2 are horizontal velocity errors; they are directly measured:

$$z_i(k) = \delta V_i(k) + r_i(k), \qquad i = 1, 2,$$

where $r_i(k)$ are measurement noises; β_1, β_2 are kinematic horizontal errors; κ_i were defined above. Thus, $n = 8$, $l = 4$, and $m = 2$. The detailed description of the initial differential equations for continuous signals and the passage to discrete system is presented in [4]. This description takes more than two pages and therefore is omitted here. The estimation algorithm filters out high-frequency noises of sensors and, as a result, we obtain the estimates of biases.

B. Results

The numerical experiment consisted of processing the measurements that were simulated on the basis of detailed numerical model of the phenomenon with the sampling step of 1 s. The (stepwise) jumps in the biases of the readings of one accelerometer and one gyroscope were modeled. This is equivalent to outliers in the appropriate entries of q_i. The duration of measurement process was equal to 30 min. A series of outliers was also added in the measurement noise. Two methods for numerical solution of l_1-norm approximation problem were compared: the reduction to an LP [4] and the method of weight and time recurrences from Section II. The use of LP at such volume of measurements requires a considerable computational resources. Hence let us first study an example with the step of discrete dynamic model and the step in measurements equal to 10 s, i.e., the initial set of measurement is thinned out.

The plots below show the results of estimating a part of state vector (δV_2 and κ_1) by means of three approaches: (a) l_1-norm approximation based on LP (dashed line); (b) l_1-norm approximation by means of weight and time recurrences (dash-dot line); (c) standard l_2-norm approximation without varying the weight coefficients (dotted line).

The estimates of velocity error δV_2 are depicted in Fig. 1. The 'true measurement' signal $z_2(k)$ with outliers is depicted by a solid line. Two l_1-norm estimates and the 'true' signal z_2 almost merge. Only l_2-norm estimate deviates from z_2 because of the presence of outliers. As is well known, the least squares method is sensitive to outliers in measurements in contrast to the least absolute deviations method.

The estimates of accelerometer error κ_1 are presented in

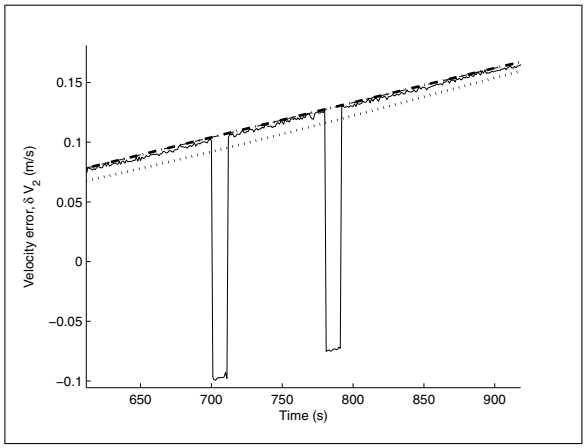

Fig. 1. Outliers in the measurements z_2 for $\delta V_2\,(m/s)$

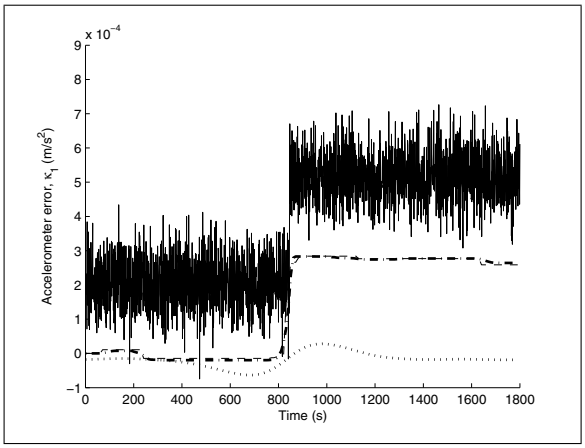

Fig. 2. Estimates for accelerometer error $\kappa_1\,(m/s^2)$

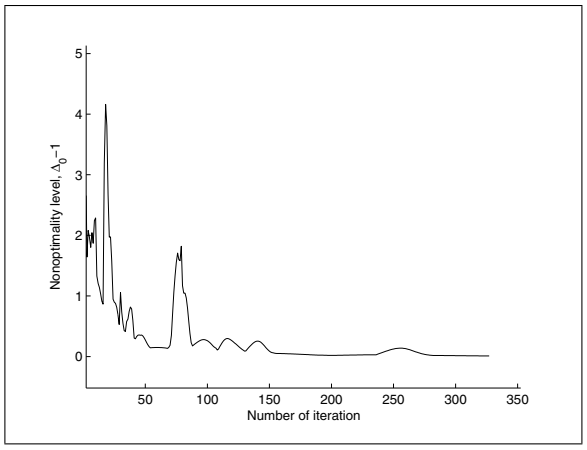

Fig. 3. Normalized guaranteed nonoptimality level $\Delta_0 - 1$

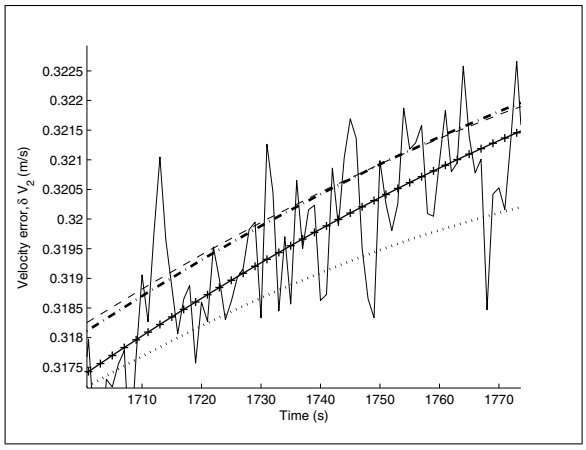

Fig. 4. Improvement in accuracy due to the decrease of step

Fig. 2. The 'true' accelerometer errors are shown by a solid line. We see that the l_1-norm based estimates of sensor bias well reflect the fact of jump and allow us to evaluate its magnitude and the instant when it occurred. The standard l_2-norm approximation gives noticeably worse results. Note that the dynamic system observability properties cause the shift in the estimates of bias. However this does not prevent the correct identification of jump. The results for the estimates of gyroscope biases are similar.

The (normalized) guaranteed nonoptimality level $\Delta_0 - 1$ is presented in Fig. 3. At the last iteration (with $\epsilon = 10^{-4}$) the accuracy was high: $\Delta_0 = 1+1.39\cdot 10^{-3}$ and the solution was found in 335 iterations over s.

A large-scale fragment of the graph for the estimates of δV_2, which were previously indicated in Fig.1, is presented in Fig. 4. It follows from Fig. 4 that a sufficiently large sampling step ($10\,s$) leads to a noticeable accumulation of divergence in estimating δV_2.

The accuracy can be improved by using the diminished sampling step in the algorithm. The smaller is the sampling step, the greater amount of measurements are taken into account, and the smaller are the errors of discretization. For example, one can take the step equal to 1 second. However, with this step, the dimension of vectors and matrices in the corresponding static problem becomes too large and the LP-based algorithms stop working on personal computer. When employing Weiszfeld's algorithm with the reduction to static problem, the difficulties with memory do not arise, but the computational process lasts for hours. The algorithm with weight and time recurrences, which was described in Section II (Weiszfeld's algorithm for dynamic case), solves the problem in a matter of minutes. The l_1-estimate of δV_2 obtained by Weiszfeld's algorithm from Section II for the step equal to $1\,s$ is displayed in Fig. 4 by a solid line with crosses. We see that the systematic discrepancy between the estimate of δV_2 and the corresponding measured signal z_2 was eliminated. The solution was obtained in 398 iterations over s with $\Delta_0 = 1 + 7.52 \cdot 10^{-4}$.

C. Discussion

Thus a tenfold increase in the number of measurements did not lead to a significant increase in the number of iterations in the outer loop. Therein lies the advantage of the method of weight and time recurrences proposed in the paper: instead of solving a variational problem with a large number of variables, a sequence of operations is carried out with vectors and matrices that are comparable in dimension with the state vector of dynamic system under consideration. This allows us to process large arrays of measurements.

V. CONCLUSION

The method of weight and time recurrences has demonstrated its efficiency: the numerical experiments and the values of nonoptimality level show that the algorithm's iterations are successful. Moreover, the method of weight and time recurrences has some advantages over more customary methods in ease of computation and working storage saving.

ACKNOWLEDGMENT

The authors thank the reviewers for helpful comments.

REFERENCES

[1] P. Bloomfield and S. L. Steiger, Least Absolute Deviations: Theory, Applications, and Algorithms. Boston-Basel-Stuttgart: Birkhaüser, 1983.

[2] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge: Cambridge University Press, 2004.

[3] V. I. Mudrov and V. L. Kushko, Methods for Processing the Observations: Quasi-Likelihood Estimates. Moscow: Radio and Communication, 1983 (in Russian).

[4] P. A. Akimov and A. I. Matasov, State estimation via l_1-norm approximation: application to inertial navigation, in Proc. 18th IFAC World Congress, Milano, 2011, pp. 14386-14391, Paper No. 0123.

[5] P. Akimov and A. Matasov, Estimation of biases in gyro drifts by means of l_1-norm approximation, in Proc. 7th IFAC Symposium on Robust Control Design, Aalborg, 2012, pp. 436-441.

[6] T. Kailath, A. H. Sayed, and B. Hassibi, Linear Estimation. New Jersey: Prentice Hall, 2000.

[7] A. I. Matasov, Estimators for Uncertain Dynamic Systems. Dordrecht-Boston-London: Kluwer, 1999.

[8] I. Ekeland and R. Temam, Convex Analysis and Variational Problems. Amsterdam-Oxford: North-Holland Publ. Co. and New York: American Elsevier Publ. Co., 1976.