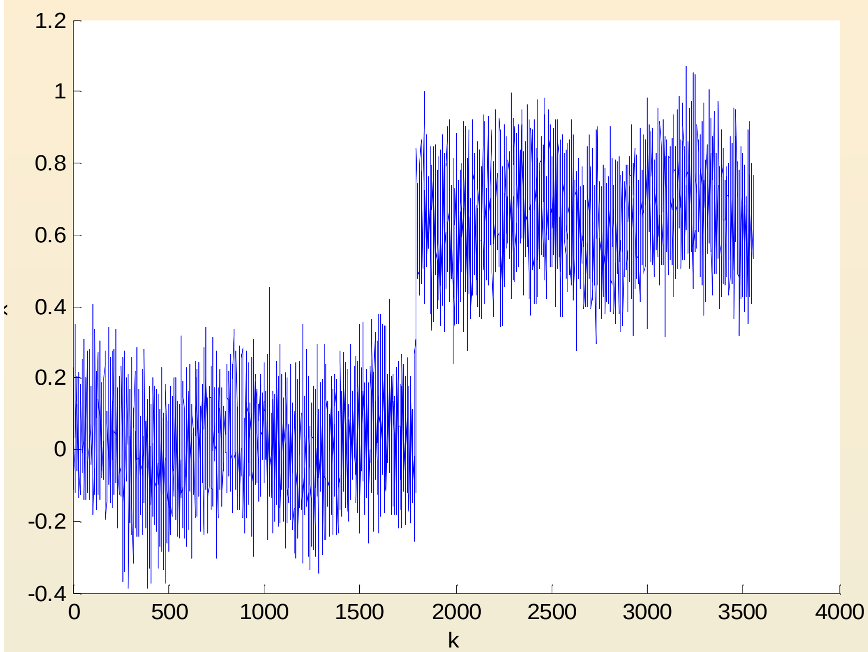
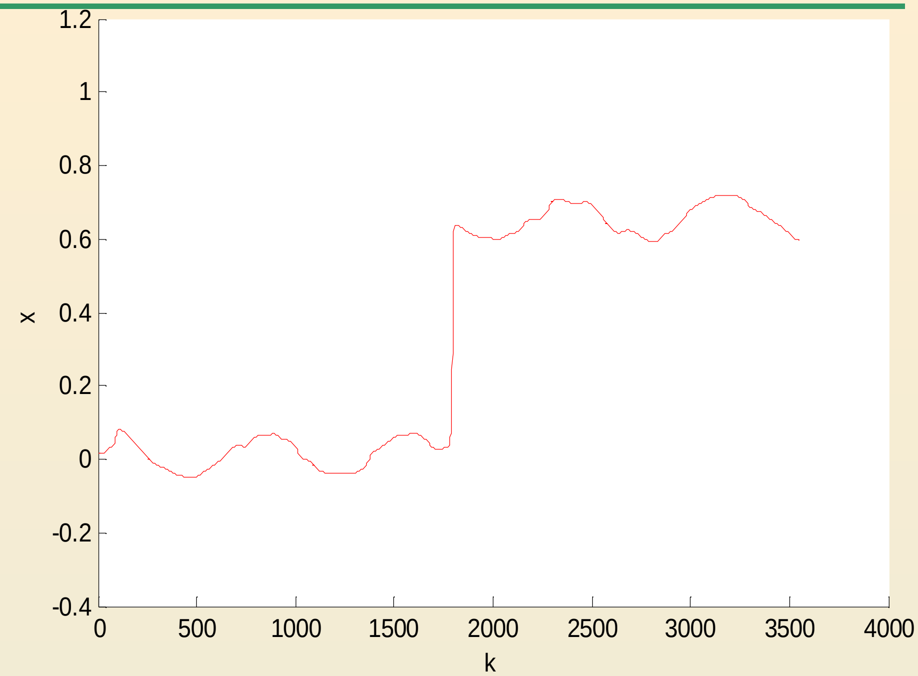
DETECTION OF JUMPS BY MEANS OF l_1-NORM APPROXIMATION

Akimov P.A., Matasov A.I.

Jumps in dynamic systems





Stepwise change (jump)

$$x(t) = \begin{cases} x_1(t) & \text{if} \quad t \in [0, T_0] \\ x_2(t) & \text{if} \quad t \in (T_0, T] \end{cases}$$

Simple differential model

$$\dot{x}(t) = q_1(t) + (x_2 - x_1)\delta(t - T_0)$$

Applications:

- Target tracking
- Navigation (sensor calibration)
- Control problems in Mechanics, Finance, Biology, Telecommunication
- Image Processing

One-dimensional continuous time system

The estimation of the state $x(t) \in \mathbf{R}$

$$\dot{x}(t) = q(t) \ , \qquad t \in [0,T]$$

$$z(t) = x(t) + \rho(t) \ , \qquad t \in [0,T]$$

Noises: $q(t), \rho(t)$ Weight coefficients: $Q, R > 0$

Least Squares Method (LSM) , $L_2 -$ norm approximation

$$\int_o^T R^{-2}(z(t) - x(t))^2 dt + \int_o^T Q^{-2}\dot{x}^2(t)dt \to \inf_{x(\cdot)\in W_{2,1}[0,T]}$$

Least Absolute Deviations Method (LADM) , $L_1 -$ norm approximation

$$\int_o^T R^{-1} \,|\, z(t) - x(t) \,|\, dt + \int_o^T Q^{-1} \,|\, \dot{x}(t) \,|\, dt \to \inf_{x(\cdot)\in W_{1,1}[0,T]}$$

The problem is to estimate the stepwise jumps in $x(t)$.

Solutions of the continuous time estimation problems

Lemma. *The LSM -solution is described by the boundary-value problem*

$$\ddot{x}(t) = \omega^2 (x(t) - z(t)), \qquad \dot{x}(0) = 0, \quad \dot{x}(T) = 0, \qquad (\omega = Q / R)$$

Explicit solution

$$x(t) = \frac{\omega \operatorname{ch}(\omega t)}{\operatorname{sh}(\omega T)} \int_0^T \operatorname{ch}(\omega(T - s)) z(s) ds + \omega \int_0^t \operatorname{sh}(\omega(t - s)) z(s) ds.$$

Let $z(t)$ be a stepwise function $(\rho(t) = 0)$

$$z(t) = \begin{cases} 0 \text{ if } t \in [0, T_0] \\ 1 \text{ if } t \in (T_0, T] \end{cases}$$

Lemma. *Suppose* $\omega(T - T_0) > 1, \quad \omega T_0 > 1;$ *then the LADM-solution has the form*

$$x(t) = z(t).$$

In contrast to LSM , the LADM transmits the input signal without distortion.

Discrete dynamic system

$$X(k+1) = FX(k) + Gq(k) + g(k), \qquad k = 0,\dots,K-1$$

$$X(k) \in \mathbf{R}^n, \quad q(k) \in \mathbf{R}^l$$

Measurements

$$z(k) = H\,X(k) + r(k), \qquad z(k), r(k) \in \mathbf{R}^m, \qquad k = 0,\dots,K$$

A priori data

$$\tilde{X}(0) = X(0) + \tilde{r}(0), \qquad \tilde{X}(0), \tilde{r}(0) \in \mathbf{R}^n$$

Typical magnitudes of noises

$$\tilde{r}_i(0) \sim \Pi_i, \qquad i = 1,\dots,n$$

$$r_i(k) \sim R_i, \qquad i = 1,\dots,m$$

$$q_i(k) \sim Q_i, \qquad i = 1,\dots,l$$

Weight matrices

$$\Pi = \text{diag}\,\{\Pi_1,\dots,\Pi_n\}$$

$$R = \text{diag}\,\{R_1,\dots,R_m\}$$

$$Q = \text{diag}\,\{Q_1,\dots,Q_l\}$$

Optimization problem

l_1 – norm approximation problem

$$I(X,q) = \| \Pi^{-1}(\tilde{X}(0) - X(0)) \|_1 + \sum_{k=0}^{K} \| R^{-1}(z(k) - H\ X(k)) \|_1 + \sum_{k=0}^{K-1} \| Q^{-1} q(k) \|_1 \to \min_{X,\ q}$$

subject to

$$X(k+1) = FX(k) + Gq(k) + g(k), \qquad k = 0,\dots,K-1 \qquad (1)$$

Stochastic framework: if $\tilde{r}_i(0),\ r_i(k),\ q_i(k)$ have the Laplace distribution density function, then (1) is equivalent to the maximization of an a posteriori density function.

Numerical methods:

- reduction to LP
- Alternating Direction Method of Multipliers (ADMM)
- quadratic approximation of nonsmooth functions

Weight and Time Recursions (I)

Basic idea: quadratic approximations of the nonsmooth functions

$$|z_j(k) - H_j X(k)| \approx \frac{(z_j(k) - H_j X(k))^2}{|z_j(k) - H_j X(k,s)|}$$

$(X^{(s)},\ q^{(s)}) = \{X(0,s),\ldots,X(K,s),\ q(0,s),\ldots,q(K-1,s)\}$ a solution from the previous step

Two nested iteration processes

«Weight» recursions: a sequence of quadratic optimization problems

$$J_{(s+1)} = \min_{X,q} (\tilde{X}(0) - X(0))^T \Pi_W^{-2}(s)(\tilde{X}(0) - X(0)) + \sum_{k=0}^{K-1} q^T(k) Q_W^{-2}(k,s) q(k) +$$

$$\sum_{k=0}^{K} (z(k) - H\ X(k))^T R_W^{-2}(k,s)\,(z(k) - H\ X(k) = \min_{X,q} J(X,q,s) \qquad (3)$$

subject to $X(k+1) = FX(k) + Gq(k) + g(k), \quad k = 0,\ldots, K-1$

A regularization is performed to avoid small residuals

$$|z_j(k) - H_j X(k,s)| \approx \alpha R_j$$

Weight and Time Recursions (II)

Weight matrices

$$\Pi_{Wj}^{-2}(s) = \begin{cases} \dfrac{\Pi_j^{-1}}{|\tilde{X}_j(0) - X_j(0,s)|} & \text{if } |\tilde{X}_j(0) - X_j(0,s)| > \alpha\Pi_j \\ \Pi_j^{-2}/\alpha & \text{if } |\tilde{X}_j(0) - X_j(0,s)| \le \alpha\Pi_j \end{cases}$$

$$Q_{Wj}^{-2}(k,s) = \begin{cases} \dfrac{Q_j^{-1}}{|q_j(k,s)|} & \text{if } |q_j(k,s)| > \alpha Q_j \\ Q_j^{-2}/\alpha & \text{if } |q_j(k,s)| \le \alpha Q_j \end{cases}$$

$$R_{Wj}^{-2}(k,s) = \begin{cases} \dfrac{R_j^{-1}}{|z_j(k) - H_j X(k,s)|} & \text{if } |z_j(k) - H_j X(k,s)| > \alpha R_j \\ R_j^{-2}/\alpha & \text{if } |z_j(k) - H_j X(k,s)| \le \alpha R_j \end{cases}$$

$$(X^{(s+1)},\ q^{(s+1)}) = \arg\min_{X,q} J(X,q,s)$$

subject to

$$X(k+1) = FX(k) + Gq(k) + g(k),$$

$$k = 0,\ldots, K-1$$

The sequence of steps $\{X^{(s)},\ q^{(s)}\}_{s=0,1,2,\ldots}$

allows to approximate the exact solution

Weight and Time Recursions (III)

«Time» recursions: the solution of quadratic smoothing problem via Bryson-Frazier formulas

$$X(k,s+1) = X^-(k) + P^-(k)\lambda(k), \qquad k = 0,\ldots,K,$$
$$q(k,s+1) = Q_W^2(k,s)\, G^T \lambda(k+1), \qquad k = 0,\ldots,K-1,$$

$X^-(k),\ P^-(k)$ are obtained by the KF

$$X^-(k+1) = FX^-(k) + K_p(k)(z(k) - HX^-(k)) + g(k), \qquad X^-(0) = \tilde{X}(0),$$
$$P^-(k+1) = FP^-(k)F^T + G\,Q_W^2(k,s)G^T$$
$$- K_p(k)\ (R_W^2(k,s) + HP^-(k)H^T)K_p^T(k), \quad P^-(0) = \Pi_W^2(k,s),$$
$$K_p(k) = FP^-(k)H^T(R_W^2(k,s) + HP^-(k)H^T)^{-1}, \quad k = 0,\ldots,K-1,$$

backwards-time recursions for $\lambda(k)$

$$\lambda(k) = (F - K_p(k)H)^T \lambda(k+1)$$
$$+ H^T(R_W^2(k,s) + HP^-(k)H^T)^{-1}(z(k) - HX^-(k)),$$
$$k = 0,\ldots,K, \quad \lambda(K+1) = 0$$

Nonoptimality levels

Nonoptimality level of a current iteration

$$\Delta = \frac{I(X^{(s+1)},\ q^{(s+1)})}{I_0}$$

$$\Delta \geq 1$$

I_0 is the value of problem (1)

Guaranteed nonoptimality level $\Delta \leq \Delta_0$ $\Delta_0 = ?$

Stopping criterion for the algorithm of weight and time recursions $\Delta_0 \leq \Delta_{\text{end}}$

Theorem. *Let* $(X^{(s+1)}, q^{(s+1)})$ *be the solution of the quadratic approximation problem (2). Then the guaranteed nonoptimality level for the* $l_1 -$ *norm approximation problem is determined by the formula*

$$\Delta \leq \Delta_0, \quad \Delta_0 = \frac{I(X^{(s+1)},\ q^{(s+1)}) \cdot \theta}{J^2_{(s+1)}}$$

The duality theory of convex variational problems is used

$$\theta = \max\{\| \Pi\ \Pi_W^{-2}(s)\,(\tilde{X}(0) - X(0, s+1))\|_\infty,$$
$$\| R\ R_W^{-2}(k,s)\,(z(k) - HX(k, s+1))\|_\infty,$$
$$\| Q\ Q_W^{-2}(i,s)\, q(i, s+1)\|_\infty\}_{i=0,\dots,K-1}^{k=0,\dots,K}$$

Mixed-norm approximation

l_1/l_2–norm approximation problem

$$\Phi(X,q)=\|\Pi^{-1}(\tilde{X}(0)-X(0))\|_2^2+\sum_{k=0}^{K}\|R^{-1}(z(k)-H\ X(k))\|_2^2+\sum_{k=0}^{K-1}\|Q^{-1}q(k)\|_1\to\min_{X,\ q}$$

subject to $X(k+1)=FX(k)+Gq(k)+g(k),\quad k=0,\dots,K-1$ (3)

l_1 terms correspond to the variables that can admit the jumps

The algorithm of weight and time recursions is applied to problem (3). Only $Q_W^{-2}(i,s)$ matrices are updated at the “weight variation” steps

Theorem. *Let* $(X^{(s+1)},\ q^{(s+1)})$ *be the solution of the quadratic approximation problem. Then the guaranteed nonoptimality level for the mixed* l_1/l_2- *norm approximation problem is determined by the formula*

$$\Delta\le\Delta_0,\quad \Delta_0=\Phi(X^{(s+1)},\ q^{(s+1)})\left(-\theta_2\min\left\{\frac{J_{(s+1)}}{\theta_2},\frac{1}{\theta_\infty}\right\}^2+2J_{(s+1)}\min\left\{\frac{J_{(s+1)}}{\theta_2},\frac{1}{\theta_\infty}\right\}\right)^{-1}$$

$$\theta_\infty=\max\{\ \|Q\,Q_W^{-2}(i,s)\,q(i,s+1)\|_\infty\}_{i=0,\dots,K-1}$$

$$\theta_2=\|\Pi^{-1}(\tilde{X}(0)-X(0,s+1))\|_2^2+\sum_{k=0}^{K}\|R^{-1}(z(k)-H\ X(k,s+1))\|_2^2$$

Numerical example

Discrete dynamic system

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(k+1) = \begin{pmatrix} 1 & 0.04 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(k) + \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}(k), \qquad k = 1,\dots, K-1$$

Measurements and a priori data

$$z(k) = x_1(k) + r(k), \quad k = 0,\dots, K, \qquad \tilde{X}(0) = (0,\ 0)^T$$

Weight matrices

$$\Pi = \operatorname{diag}(10,\ 1), \quad Q = \operatorname{diag}(0.2,\ 0.2), \quad R = 3$$

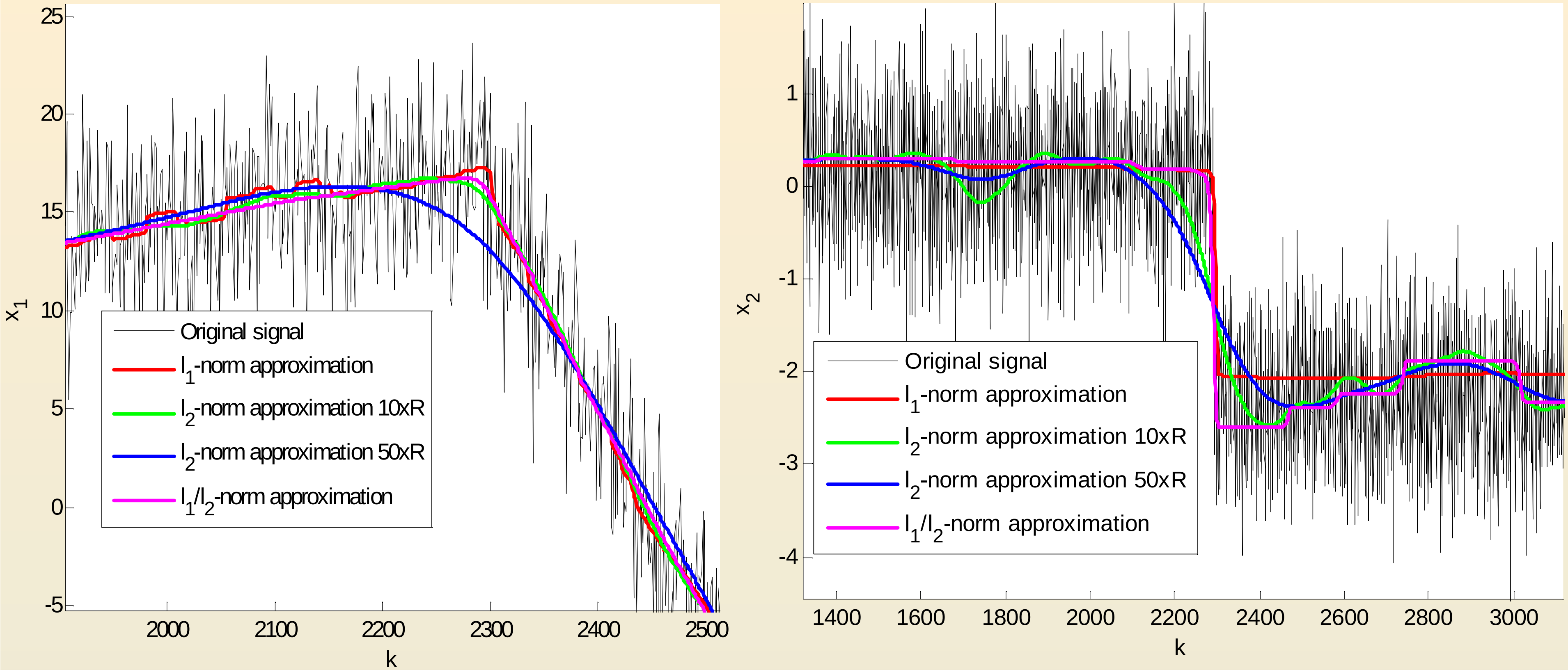
K=3600

14389 variables

10800 summands in the cost function

A stepwise jump in $x_2(k)$

Estimates



Numerical methods

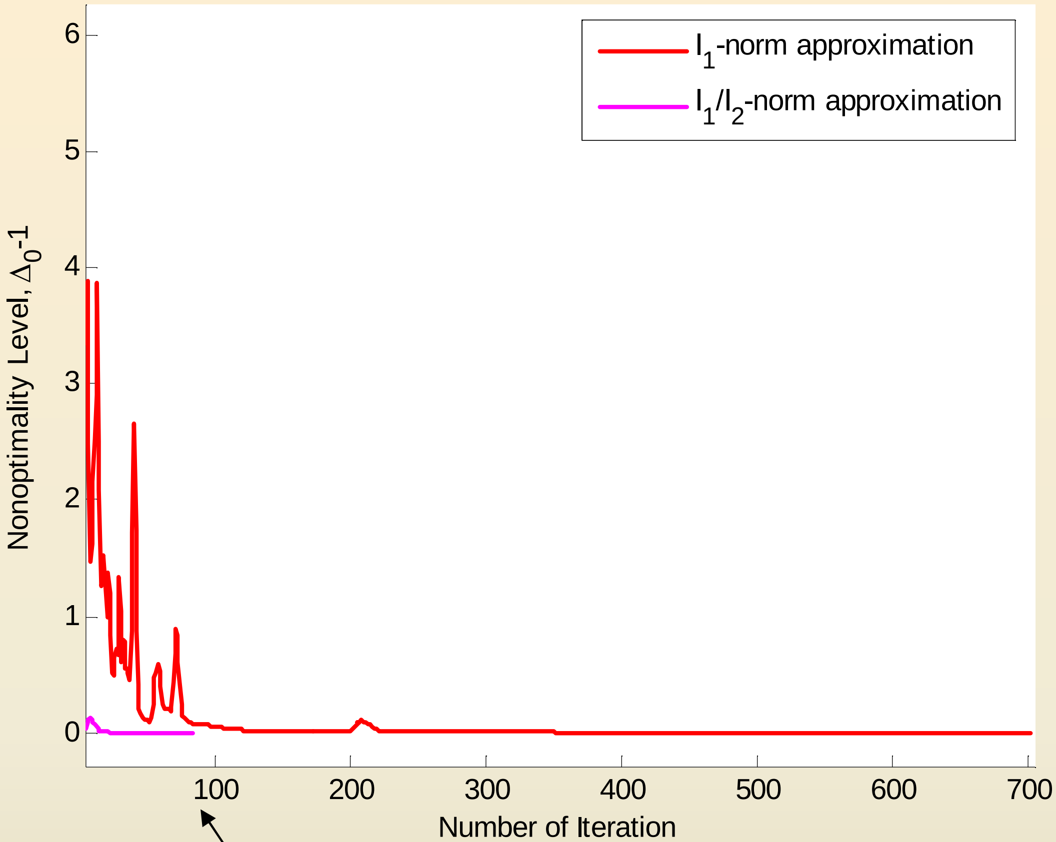
$l_1 -$ norm approximation, weight and time recursions: the solution was found in 943 *s*

$l_1/l_2 -$ norm approximation, weight and time recursions: the solution was found in 80 *s*

$l_2 -$ norm approximation, Kalman smoother formulas, R x 10: the solution was found in 0.6 *s*

$l_2 -$ norm approximation, Kalman smoother formulas, R x 50: the solution was found in 0.6 *s*

Nonoptimality levels



Algorithm parameters

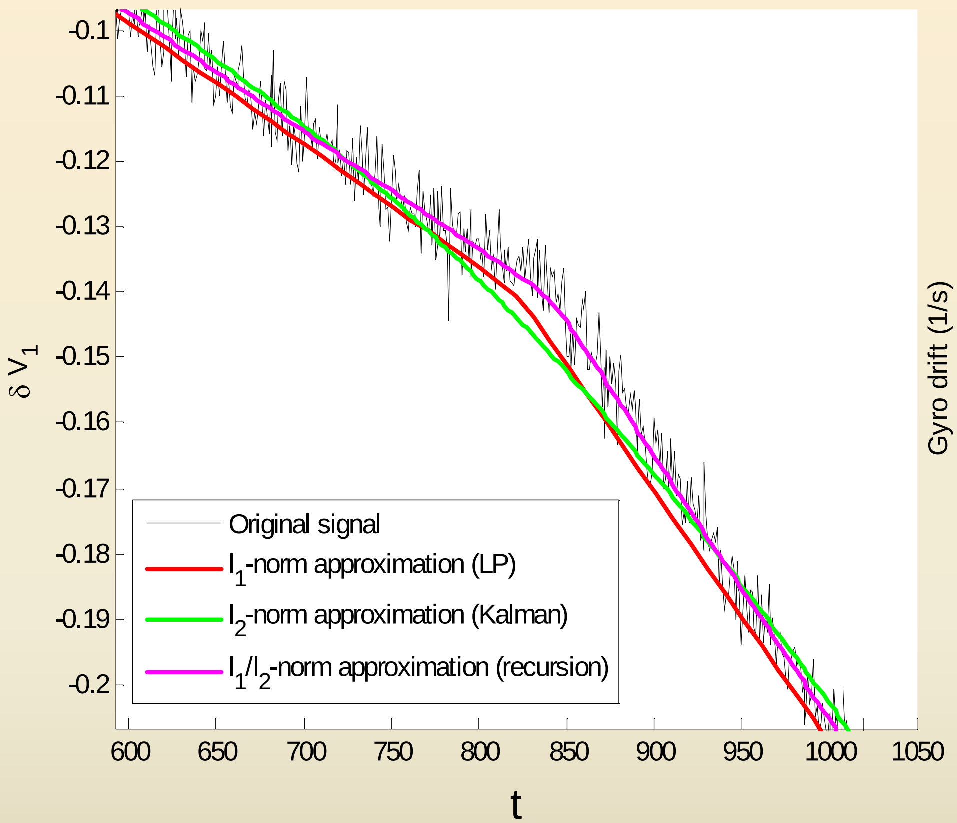
$$\alpha = 10^{-6}$$

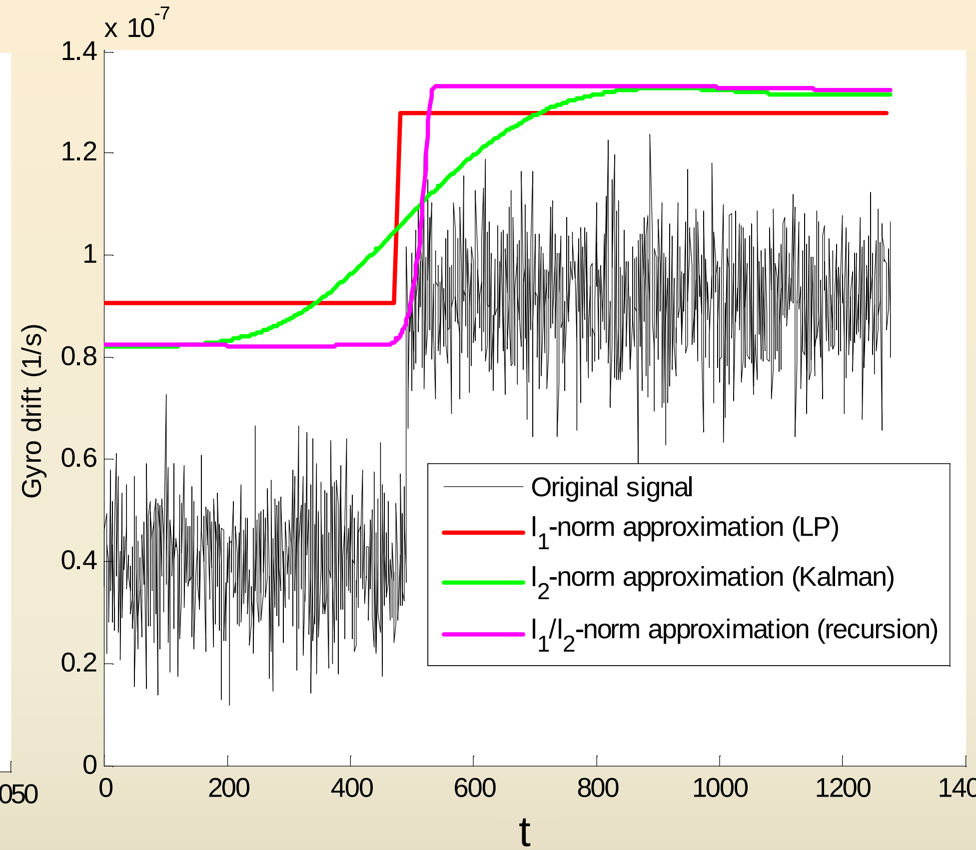
$$\Delta_{end} = 1 + 10^{-3}$$

Mixed-norm approach requires significantly less number of iterations than the l_1 – norm approximation

Application to inertial navigation

Jump detection in a gyroscope of a strapdown inertial navigation system (SDINS)





$n = 8,\ l = 4,\ m = 2,$

$K = 1280$

It is difficult to apply an LP to the problems with a large number of variables. Hence the sampling step equals 10 *s* for LP. The algorithm of weight and time recursions allows to solve the nonsmooth optimization problems (1) and (3) with the sampling step 1 *s*.

Conclusion

- The application of $l_1 -$ or $l_1 / l_2 -$ norm approximation allows to estimate effectively the state vector with possible rare jumps. This property follows from the sparsity of solutions.

- The algorithm of weight and time recursions reduces the original nonsmooth problem to a sequence of quadratic optimization problems (“weight” recursion). In case of dynamic systems such problems are solved by the recurrent Kalman smoother (“time” recursion).

- The guaranteed nonoptimality levels of current iterations characterize the accuracy of approximate solutions and give a criterion for stopping the calculation process.

- The duality theory of convex variational problems is used for constructing the guaranteed nonoptimality levels.

- The algorithm of weight and time recursions is simple for implementation and demonstrates the high efficiency in case of large measurement data.

References

[1] **B. T. Polyak, “**Methods of $l_1 -$ optimization in control and filtering,” in *Proc. 3rd Multi-Conference on Control Problems*, State Research Center of the Russian Federation, Concern CSRI Elektrobribor, JSC, St. Petersburg, Russia, Oct. 2010.

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

[3] **P. Bloomfield, W.L. Steiger,** *Least Absolute Deviations: Theory, Applications, and Algorithms.* Boston-Basel-Stuttgart: Birkhauser, 1983.

[4] **I. Ekeland, R. Temam,** *Convex Analysis and Variational Problems.* Amsterdam, New York: North Holland Pub. Co., American Elseveir Pub. Co., 1976.

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

[6] **S. Farahmand, G. Giannakis, D. Angelosante, “**Doubly robust smoothing of dynamical processes via outlier sparsity constraints,” *IEEE Trans. Signal Process.,* vol. 59, no. 10, pp. 4529–4543, 2011.

[7] **P. A. Akimov, A. V. Derevyankin, A. I. Matasov,** *Guaranteed Approach and* l_1 *–norm Approximation in Parameter Estimation Problems for SDINS under Bench Testing.* Moscow, Russia: Moscow Univ. Press, 2012 (in Russian).

[8] **P. A. Akimov, A. I. Matasov,** “Recursive estimation algorithm for $l_1 -$ norm approximation in dynamic systems with nonoptimality levels,” *Proc. European Control Conf.*, Zurich, Switzerland, pp. 2427-2432, 2013.

[9] **P. A. Akimov, A. I. Matasov,** “Weight and time recursions in dynamic state estimation problem with mixed-norm cost function,” *IEEE Trans. Automatic Control,* vol. 60, no. 4, pp.1050-1063, 2015.