RC24709 (W0812-069) December 11, 2008
Computer Science

IBM Research Report

A Simple Method for Sparse Signal Recovery from Noisy Observations Using Kalman Filtering

Avishy Carmi
Cambridge University
UK

Pini Gurfil
Technion,
Haifa, Israel

Dimitri Kanevsky
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
USA

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Abstract

We present a simple method for recovering sparse signals from a series of noisy observations. Our algorithm is a Kalman filter (KF) that utilize a so-called pseudo-measurement technique for optimizing the convex minimization problem following from the theory of compressed sensing (CS). Compared to the recently introduced CS-KF in [1] which involves the implementation of an additional CS optimization algorithm (e.g., the Dantzig selector), our method is remarkebly easy to implement as it is exclusively based on the KF formulation. The results of an extensive numerical study are provided demonstrating the performance and viability of the new method.

Sparse Signal Recovery

Consider an $\mathbb{R}^n$-valued random discrete-time process $\{x_k\}_{k=1}^{\infty}$ that is sparse in some known orthonormal sparsity basis $\psi \in \mathbb{R}^{n\times n}$, that is

$$z_k = \psi^T x_k, \quad |\text{supp}(z_k)| << n \tag{1}$$

where $\text{supp}(z_k)$ denotes the support of z_k. Assume that z_k evolves according to

$$z_{k+1} = Az_k + w_k \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$ and $\{w_k\}_{k=1}^{\infty}$ is a zero-mean white Gaussian sequence with covariance Q_k. The process x_k is measured using a sequence of noisy observations given by

$$y_k = Hx_k + \zeta_k = H'z_k + \zeta_k \tag{3}$$

where $\{\zeta_k\}_{k=1}^{\infty}$ is a zero-mean white Gaussian sequence with covariance R_k, and $H := H'\psi^T \in \mathbb{R}^{m\times n}$ with $m < n$.

Letting $y^k := [y_1, \dots, y_k]$, our problem is defined as follows. We are interested in a y^k-measurable estimator $\hat{x}_k$ such that the minimum mean square error (MMSE) $E\left[\| x_k - \hat{x}_k \|_2^2\right]$ is minimized. It is well known that if the above linear system (2),(3) is observable, i.e.,

$$\mathscr{O} := \begin{bmatrix} H \\ HA \\ \vdots \\ HA^{n-1} \end{bmatrix}, \quad \text{rank}(\mathscr{O}) = n \tag{4}$$

then the solution to this problem can be obtained using the Kalman filter (KF). On the other hand, if the system is unobservable then the regular KF algorithm is useless in this case.

When the above system is unobservable, if for instance $A = I_{n\times n}$, then it may look hopeless to try reconstructing x_k from an under-determined system in which $\text{rank}(H) < n$. Surprisingly, this problem may be circumvented by taking into account the fact that z_k is sparse.

The Combinatorial Problem and Compressed Sensing

It has already been shown that in the deterministic case (i.e., when z is a parameter vector) one can recover z (and therefore also x, i.e., $x = \psi z$) with high accuracy by solving the optimization problem [2, 3]

$$\min \| \hat{z} \|_0 \ \text{ s.t. } \sum_{i=1}^{k} \| y_i - H'\hat{z} \|_2^2 \leq \epsilon \tag{5}$$

for sufficiently small ϵ. Following a similar approach, in the stochastic case it can be shown that the sought-after optimal estimator satisfies

$$\min \| \hat{z}_k \|_0 \ \text{ s.t. } E_{z_k|y^k}\left[\| z_k - \hat{z}_k \|_2^2\right] \leq \epsilon \tag{6}$$

Unfortunately, the above optimization problems are NP-hard and cannot be solved efficiently. Recently, it has been shown that if the sensing matrix H' obeys a so-called *restricted isometry hypothesis* (RIH) then the solution of the combinatorial problem (5) can almost always be obtained by solving the convex optimization [2, 4]

$$\min \| \hat{z} \|_1 \ \text{ s.t. } \sum_{i=1}^{k} \| y_i - H'\hat{z} \|_2^2 \leq \epsilon \tag{7}$$

This is a fundamental result in a new emerging theory known as compressed sensing (CS) [2, 4]. The essential idea here is that the convex l_1 minimization problem can be efficiently solved. Another insights provided by CS are related to the construction of sensitivity matrices that satisfy the RIH. These underlying matrices are random by nature which in turn has shed a new light on the way observations should be sampled. For an extensive review of CS the reader is referred to [2, 4].

CS-Embedded Kalman Filtering

Inspired by the CS approach while retaining the KF objective function we replace (6) with

$$\min_{\hat{z}_k} E_{z_k|y^k}\left[\| z_k - \hat{z}_k \|_2^2\right] \text{ s.t. } \| \hat{z}_k \|_1 \leq \epsilon' \tag{8}$$

The above constrained optimization problem can be easily solved in the framework of Kalman filtering using a so-called pseudo measurement (PM) technique [5,6]. Following this, the inequality constraint $\| z_k \|_1 \leq \epsilon'$ is incorporated into the filtering process using a fictitious measurement

$$0 = \| z_k \|_1 - \epsilon' \tag{9}$$

where ϵ' serves as a measurement noise. The above PM can be rewritten as

$$0 = \bar{H} z_k - \epsilon', \quad \bar{H} := [\text{sign}(z_k(1)), \ldots, \text{sign}(z_k(n))] \tag{10}$$

where $\text{sign}(z_k(i))$ denotes the sign function of the ith element of z_k, that is

$$\text{sign}(z_k(i)) = \begin{cases} +1, & \text{if } z_k(i) \geq 0 \\ -1, & \text{if } z_k(i) < 0 \end{cases} \tag{11}$$

In this setting, the covariance R_ϵ of ϵ' is regarded as a tuning parameter which can be determined based on simulation runs.

It is well known in the literature that the PM stage can be iterated to better enforce the constraint [6]. Regulating the tradeoff between estimation accuracy and computational complexity, the number of PM iterations is essentially application dependent.

Algorithm Summary

A single iteration of the CS-embedded KF is detailed below.

1. *Prediction*

$$\hat{z}_{k+1|k} = A\hat{z}_{k|k} \tag{12a}$$

$$P_{k+1|k} = AP_{k|k}A^T + Q_k \tag{12b}$$

2. *Measurement Update*

$$K_k = P_{k+1|k}H'^T\left(H'P_{k+1|k}H'^T + R_k\right)^{-1} \tag{13a}$$

$$\hat{z}_{k+1|k+1} = \hat{z}_{k+1|k} + K_k\left(y_k - H'\hat{z}_{k+1|k}\right) \tag{13b}$$

$$P_{k+1|k+1} = (I - K_k H')P_{k+1|k} \tag{13c}$$

3. *CS Pseudo Measurement:* Let $P^1 = P_{k+1|k+1}$ and $\hat{z}^1 = \hat{z}_{k+1|k+1}$. For $\tau = 1, 2, \ldots, N_\tau - 1$ iterations do the following

$$\bar{H}_\tau = [\text{sign}(\hat{z}^\tau(1)), \ldots, \text{sign}(\hat{z}^\tau(n))] \tag{14a}$$

$$K^\tau = P^\tau \bar{H}_\tau^T\left(\bar{H}_\tau P^\tau \bar{H}_\tau^T + R_\epsilon\right)^{-1} \tag{14b}$$

$$\hat{z}^{\tau+1} = (I - K^\tau \bar{H}_\tau)\hat{z}^\tau \tag{14c}$$

$$P^{\tau+1} = (I - K^\tau \bar{H}_\tau)P^\tau \tag{14d}$$

4. Set $P_{k+1|k+1} = P^{N_\tau}$ and $\hat{z}_{k+1|k+1} = \hat{z}^{N_\tau}$.

Notice that this is an unusual implementation of the KF as the matrix $\bar{H}_\tau$ is random.

Numerical Study

The performance of the CS-embedded KF algorithm is demonstrated using several simple examples in which sparse signals are recovered from a series of noisy observations. Two cases are examined: 1) recovery of a static parameter vector z, and 2) recovery of a stochastic process z_k of a fixed support. In an additional example the robustness of the algorithm is examined as sudden changes occur in the process support over time.

Static Case

In this example the signal $z \in \mathbb{R}^{256}$ is assumed to be a sparse parameter vector (i.e., $A = I_{256\times256}$, $Q_k = 0$). The signal support consists of total of 10 elements $z(i) \neq 0$ of which both the index and value are uniformly sampled over $i \sim U_i[1, 256]$ [1] and $z(i) \sim U[-10, 10]$, respectively. The sensitivity matrix H' consists of 72 rows in which the elements are sampled from a Guassian distribution. The columns of H' are normalized following the example in [1] (this matrix has been shown to satisfy the RIH, see [1, 4]). The observation noise covariance is set as $R_k = 0.001^2 I_{72\times72}$ where the tuning parameter $R_\epsilon = 200^2$.

The estimation performance of the CS-embedded KF algorithm is presented in Fig. 1. This figure depicts the mean square estimation error based on $N = 50$ Monte Carlo runs for various number N_τ of PM iterations. Unsurprisingly, increasing N_τ yields an improved estimation performance as the estimation error attains lower values.

Figure 2 shows the performance of the algorithm in a single typical run.

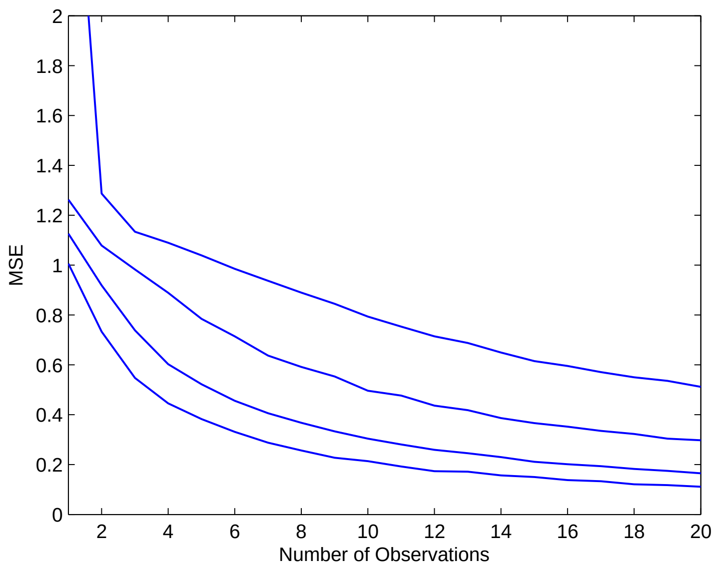

Figure 1. Mean square estimation error based on 50 runs of the CS-embedded KF for various number of PM iterations. The lines top to bottom correspond to 50, 100, 200 and 300 PM iterations. Static case.

Dynamic Case

In this example the sparse signal consists of exactly 10 non-zero elements that behave as a random walk process. The driving white noise covariance of the elements in the support of z_k is taken as $Q = 1$. This process can be described by

$$z_{k+1}(i) = \begin{cases} z_k(i) + w_k(i), \quad w_k(i) \sim \mathcal{N}(0, Q), & \text{if } z_k(i) \in \text{supp}(z_k) \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

where $i \sim U_i[1, 256]$ and $z_0(i) \sim U[-10, 10]$ as before. The sensitivity matrix H' is set as in the static case and the observation noise covariance is taken as $R_k = 0.01^2 I_{72\times72}$.

The estimation performance of the CS-embedded KF algorithm in the dynamic case is presented in Fig. 3. This figure depicts the mean square estimation error based on $N = 50$ Monte Carlo runs for various number N_τ of PM iterations. As it could be expected, the attainable estimation errors in this case are slightly higher than in the static case. Nevertheless, it seems that the algorithm manages to adequately estimate the behavior of the non-zero processes as shown in Fig. 4.

[1] $U_i[a, b]$ denotes a discrete uniform distribution of which the support are all the integers in the interval $[a, b]$.

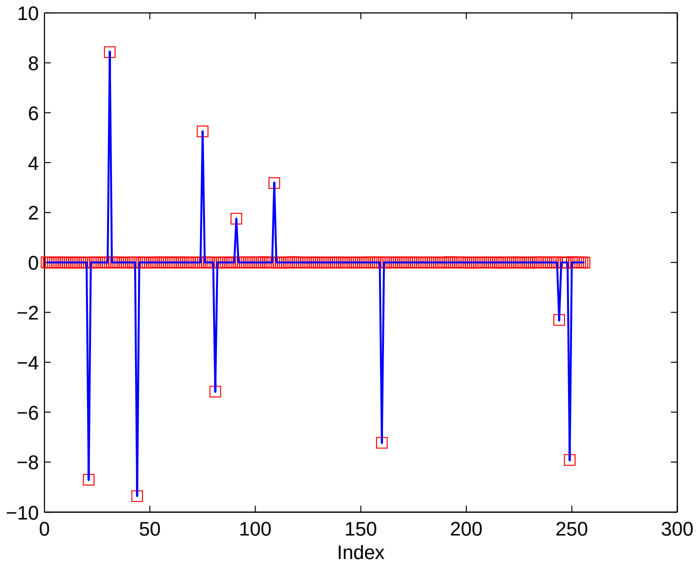

Figure 2. A typical single run of the CS-embedded KF with $N_\tau = 200$ PM iterations. Showing the elements of the true (blue line) and estimated (red squares) signals. Static case.

Sudden Change in the Support

In this example we deliberately simulate a sudden change in the support of the process z_k in (15). The change occurs at time instance $k = 10$ where 4 additional random walk processes are initiated. The indexes of these processes within z_k are uniformly sampled over the integers between 1 to 256 excluding all the indexes of the other non-zero elements.

The estimation performance of the algorithm in this case is presented in Figs. 5 and 6. From these figures it is evident that the CS-embedded KF adapts to changes in the support of the process.

Regular KF

Implementing a regular KF without the PM extension for the previous examples is useless as the underlying systems are unobservable. In the dynamic case the system is non detectable which in turn implies that the estimation error in this case may increase over time as shown in Fig. 7.

Conclusions

We have presented a simple Kalman filtering-based algorithm for sparse signal recovery from a series of noisy observations. The proposed method utilizes a so-called pseudo-measurement technique to enforce the l_1 constraint following from compressed sensing theory. The simplicity of this method originates from the fact that no considerable modifications are required in the basic Kalman filter formulation. The new algorithm is shown via simulations to adequately reconstruct sparse signals in both static and dynamic scenarios.

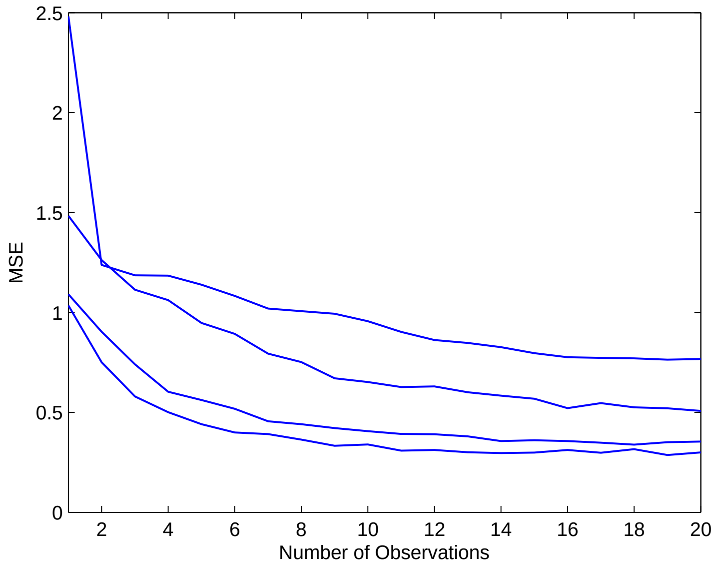

Figure 3. Mean square estimation error based on 50 runs of the CS-embedded KF for various number of PM iterations. The lines top to bottom correspond to 50, 100, 200 and 300 PM iterations. Dynamic case.

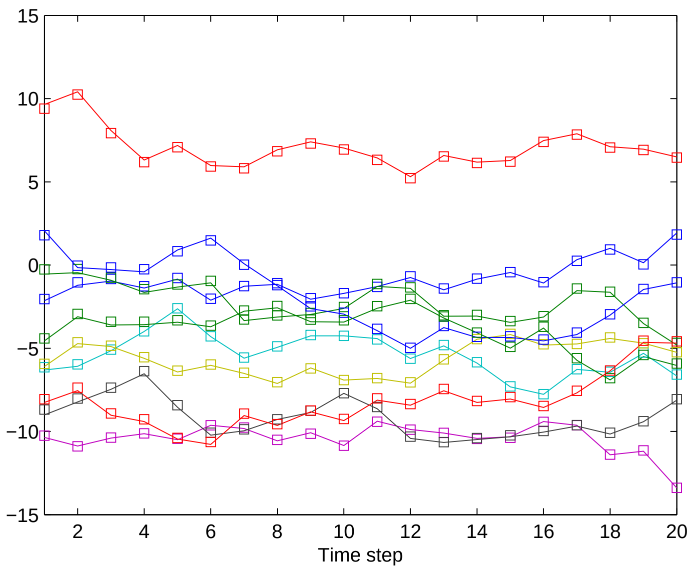

Figure 4. A typical single run of the CS-embedded KF with $N_\tau = 200$ PM iterations. Showing the true (solid lines) and estimated (squares) non-zero signals. Dynamic case.

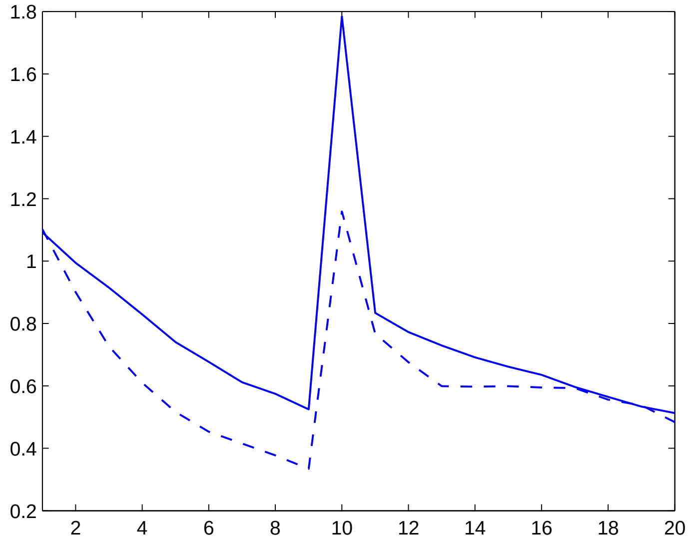

Figure 5. **Mean square estimation error based on 50 runs of the CS-embedded KF. At time $k = 10$ there is a sudden change in the support of the process. Showing the estimation performance for 100 (solid line) and 200 (dashed line) PM iterations.**

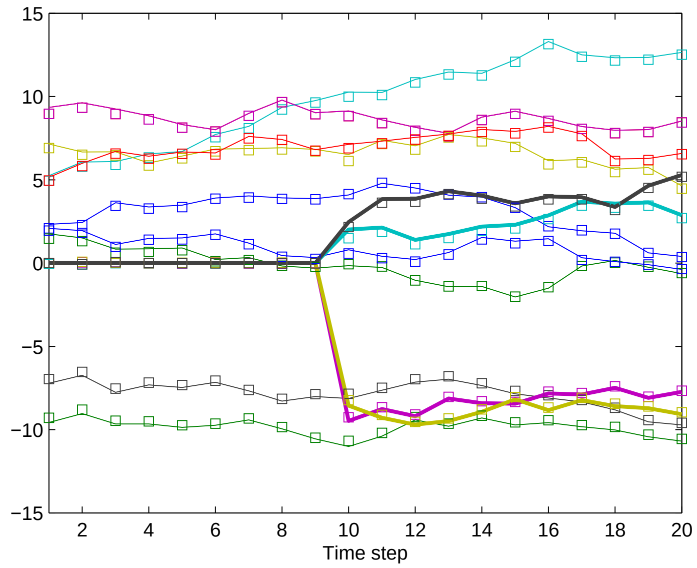

Figure 6. **A typical single run of the CS-embedded KF with $N_\tau = 200$ PM iterations. Showing the true (solid lines) and estimated (squares) non-zero signals. Bold lines represent signals that were added to the support from time $k = 10$.**

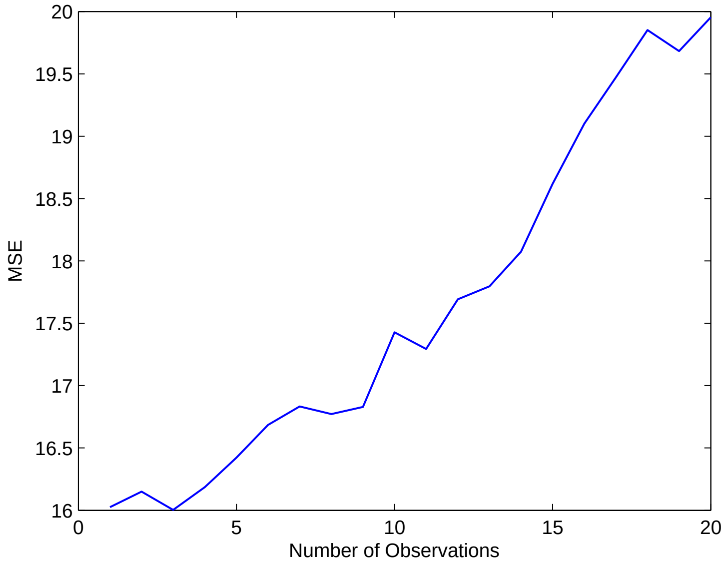

Figure 7. Mean square estimation error based on 50 runs of a regular KF (i.e., without the CS pseudo-measurement stage) applied for the dynamic signal recovery problem.

Bibliography

[1] Vaswani, N., "Kalman Filtered Compressed Sensing," Proceedings of the IEEE International Conference on Image Processing (ICIP), October 2008.

[2] Candes, E. J., Romberg, J., and Tao, T., "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Transactions on Information Theory*, Vol. 52, 2006.

[3] Chartrand, R., "Exact Reconstruction of Sparse Signals via Nonconvex Minimization," *IEEE Signal Processing Letters*, Vol. 14, 2007, pp. 707–710.

[4] Candes, E. J., "Compressive Sampling," European Mathematical Society, Proceedings of the International Congress of Mathematicians, Madrid, Spain, 2006.

[5] Deurschmann, J., Bar-Itzhack, I., and Ken, G., "Quaternion Normalization in Spacecraft Attitude Determination," American Institute of Aeronautics and Astronautics, Proceedings of the AIAA/AAS Astrodynamics Conference, 1992, pp. 27–37.

[6] Julier, S. J. and LaViola, J. J., "On Kalman Filtering With Nonlinear Equality Constraints," *IEEE Transactions on Signal Processing*, Vol. 55, No. 6, 2007, pp. 2774 – 2784.