Compressive Sensing for Gaussian Dynamic Signals

Wei Dai, Dino Sejdinovic, and Olgica Milenkovic
Emails: wei.dai1@imperial.ac.uk, d.sejdinovic@bristol.ac.uk, milenkov@uiuc.edu

Consider a linear dynamic system described by the update equations $\boldsymbol{x}_t = \boldsymbol{\Psi}_t \boldsymbol{x}_{t-1} + \boldsymbol{u}_t$, $\boldsymbol{y}_t = \boldsymbol{\Phi}_t \boldsymbol{x}_t + \boldsymbol{v}_t$. Here, $\boldsymbol{x}_t \in \mathbb{R}^n$ represents the state vector of the system, $\boldsymbol{y}_t \in \mathbb{R}^m$ denotes the measurement vector, $\boldsymbol{u}_t \in \mathbb{R}^n$ and $\boldsymbol{v}_t \in \mathbb{R}^m$ are Gaussian innovation vectors with $\boldsymbol{u}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$ and $\boldsymbol{v}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_v)$, respectively. The subscript $t = 1, 2, \cdots$ describes the time instances at which the signal is observed. Suppose that the statistics of $\boldsymbol{x}_{t-1}$ are known, and given by $\boldsymbol{x}_{t-1} \sim \mathcal{N}(\hat{\boldsymbol{x}}_{t-1}, \boldsymbol{\Sigma}_{t-1})$. Given $\boldsymbol{y}_t$, $\boldsymbol{\Psi}_t$ and $\boldsymbol{\Phi}_t$, the MAP estimate of $\boldsymbol{x}_t$, denoted by $\hat{\boldsymbol{x}}_t$, coincides with the corresponding MMSE estimate.

Now suppose that one has prior information that $\boldsymbol{x}_t$ is K-sparse. The MAP estimator that takes the sparsity assumption into consideration is given by

$$\hat{\boldsymbol{x}}_t = \underset{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0 \le K}{\arg\max}\ p_{\boldsymbol{X}_t|\boldsymbol{Y}_t, \boldsymbol{X}_{t-1}}(\boldsymbol{x}|\boldsymbol{y}_t, \hat{\boldsymbol{x}}_{t-1}),$$

where the pseudo-norm $\|\cdot\|_0$ counts the number of non-zero entries of its argument. Let $\boldsymbol{A}_t = 2\left(\boldsymbol{\Sigma}_u^{-1} + \boldsymbol{\Phi}_t^T \boldsymbol{\Sigma}_v^{-1} \boldsymbol{\Phi}_t\right)$, $\boldsymbol{b}_t = -2\left(\boldsymbol{\Sigma}_u^{-1} \boldsymbol{\Psi}_t \hat{\boldsymbol{x}}_{t-1} + \boldsymbol{\Phi}_t^T \boldsymbol{\Sigma}_v^{-1} \boldsymbol{y}_t\right)$, and $f_t(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T \boldsymbol{A}_t \boldsymbol{x} + \boldsymbol{b}_t^T \boldsymbol{x}$. It can be verified that the sparse MAP estimator is equivalent to

$$\hat{\boldsymbol{x}}_t = \underset{\boldsymbol{x}:\ \|\boldsymbol{x}\|_0 \le K}{\arg\min}\ f_t(\boldsymbol{x}). \tag{1}$$

At the first glance, the Gaussian sparse modelling looks arbitrary. The common strategy for dynamic CS usually involves certain sparsity-promoting distributions, which often result in reconstructions with high computational complexity and week performance guarantees. Note that Gaussian modelling has been successfully applied to dynamic signal processing, and that in many applications, e.g. MRI imaging, the dynamic signal at each time instance is sparse. Our model combines the advantages of both Gaussian and sparse modelling and renders good performance guarantees.

It is NP-hard to solve the optimization problem (1). We therefore propose a practical greedy algorithm to solve (1). It is based on the well-known subspace pursuit (SP) algorithm for standard compressive sensing, and therefore termed SP-MAP. The details are described in Algorithm 1. *It can be proved* that the proposed SP-MAP algorithm coincides the standard SP algorithm when $\boldsymbol{\Sigma}_u = \sigma_u^2 \boldsymbol{I}$, $\boldsymbol{\Sigma}_v = \boldsymbol{I}$ and $\sigma_u^2 \to \infty$. The *performance guarantees* of the proposed SP-MAP algorithm are based on RIP like conditions and will be detailed in the full version of this abstract.

We performed extensive numerical simulations to test our approach for K-sparse dynamical signals. In order to generate a sparse Gaussian dynamic signal, we use the model $\boldsymbol{x}_t = \mathcal{T}_k(\boldsymbol{\Phi}_t \boldsymbol{x}_{t-1} + \boldsymbol{u}_t)$, where the nonlinear mapping $\mathcal{T}_K(\boldsymbol{x})$ produces a vector that agrees with $\boldsymbol{x}$ in the K largest magnitude entries, and has all other coordinates equal to zero.

Algorithm 1 The SP-MAP Algorithm

Let $\ell_{\max}$ be the maximum iterations at each time instance. Let $\hat{\boldsymbol{x}}_0 = \boldsymbol{0}$. At time instance t, perform the following operations.

Initialization:

1) Define $\boldsymbol{x}_t' = \boldsymbol{\Psi}_t \hat{\boldsymbol{x}}_{t-1}$, $\boldsymbol{A} = 2\left(\boldsymbol{\Sigma}_u^{-1} + \boldsymbol{\Phi}_t^T \boldsymbol{\Sigma}_v^{-1} \boldsymbol{\Phi}_t\right)$ and $\boldsymbol{b} = -2\left(\boldsymbol{\Sigma}_u^{-1} \boldsymbol{x}_t' + \boldsymbol{\Phi}_t^T \boldsymbol{\Sigma}_v^{-1} \boldsymbol{y}_t\right)$.
2) Let $\ell = 0$. Let $\hat{\boldsymbol{x}}_t = -\boldsymbol{A}^{-1}\boldsymbol{b}$. Let $\mathcal{K}$ be the set of the K indices corresponding to the largest $\boldsymbol{A}_{i,i} |\hat{\boldsymbol{x}}_{t,i}|^2$'s, $i \in [n]$. Define $\hat{\boldsymbol{x}}_t^{(\ell)}$ such that $\hat{\boldsymbol{x}}_{t,\mathcal{K}^c}^{(\ell)} = \boldsymbol{0}$ and $\hat{\boldsymbol{x}}_{t,\mathcal{K}}^{(\ell)} = -\boldsymbol{A}_{\mathcal{K},\mathcal{K}}^{-1} \boldsymbol{b}_{\mathcal{K}}$.
3) Let $\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_t^{(\ell)}$. Compute $f^{(\ell)} = \frac{1}{2}\hat{\boldsymbol{x}}_t^T \boldsymbol{A} \hat{\boldsymbol{x}}_t + \boldsymbol{b}^T \hat{\boldsymbol{x}}_t$.

Iterations:

1) Let $\ell = \ell + 1$.
2) For every $i \notin \mathcal{K}$, compute $\Delta_i = \left(\langle \hat{\boldsymbol{x}}_{t,\mathcal{K}}, \boldsymbol{A}_{\mathcal{K},i}\rangle + \boldsymbol{b}_i\right)^2 / \boldsymbol{A}_{i,i}$. Let $\mathcal{K}_\Delta$ be the set of the K indices corresponding to the largest Δ_i's, $i \in \mathcal{K}^c$.
3) Let $\tilde{\mathcal{K}} = \mathcal{K} \bigcup \mathcal{K}_\Delta$. Define $\tilde{\boldsymbol{x}}_t$ such that $\tilde{\boldsymbol{x}}_{t,\tilde{\mathcal{K}}^c} = \boldsymbol{0}$ and $\tilde{\boldsymbol{x}}_{t,\tilde{\mathcal{K}}} = -\boldsymbol{A}_{\tilde{\mathcal{K}},\tilde{\mathcal{K}}}^{-1} \boldsymbol{b}_{\tilde{\mathcal{K}}}$. For every $i \in \tilde{\mathcal{K}}$, compute $\Delta_i = \boldsymbol{A}_{i,i} \tilde{\boldsymbol{x}}_{t,i}^2$.
4) Let $\mathcal{K}$ be the set of the K indices corresponding to the largest Δ_i's, $i \in \tilde{\mathcal{K}}$. Define $\hat{\boldsymbol{x}}_t^{(\ell)}$ such that $\hat{\boldsymbol{x}}_{t,\mathcal{K}^c}^{(\ell)} = \boldsymbol{0}$ and $\hat{\boldsymbol{x}}_{t,\mathcal{K}}^{(\ell)} = -\boldsymbol{A}_{\mathcal{K},\mathcal{K}}^{-1} \boldsymbol{b}_{\mathcal{K}}$. Compute $f^{(\ell)} = \frac{1}{2}\hat{\boldsymbol{x}}_t^T \boldsymbol{A} \hat{\boldsymbol{x}}_t + \boldsymbol{b}^T \hat{\boldsymbol{x}}_t$.
5) If $f^{(\ell)} > f^{(\ell-1)}$, quit the iterations.
6) Let $\hat{\boldsymbol{x}}_t = \hat{\boldsymbol{x}}_t^{(\ell)}$. If $\ell \ge \ell_{\max}$, quit the iterations. Otherwise, go to Step 1 for the next iteration.

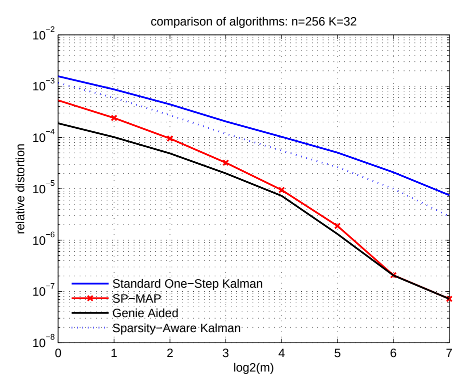

Figure 1. Comparison of reconstruction algorithms.

Figure 1 compares the proposed SP-MAP algorithm with other algorithms designed for dynamic CS. According to the simulation results, the SP-MAP algorithm outperforms others and it performs very close to the genie-aided approach when the number of samples per time instance is sufficient.

References are omitted due to the space limitation.