Journal Watch:

IEEE Transanctions Signal Processing, 15 February, 2018

Saurabh Khanna,
Signal Processing for Communication Lab, ECE, IISc

Sparse Signal Recovery Using Iterative Proximal Projection

- Sparse signal recovery problem:

$$\min_{\mathbf{x}} J(\mathbf{x}) \quad s.t. \, ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2 \leq \epsilon$$

  J is non-smooth sparsity promoting function, e.g., ℓ_0-norm, ℓ_1-norm.
- Existing liteature has focussed on J being convex.
- This work focusses on J being nonconvex & non-smooth.
- Approach: use proximal algorithms.

Sparse Signal Recovery Using Iterative Proximal Projection

- Sparse signal recovery problem:

$$\min_{\mathbf{x}} J(\mathbf{x}) \quad s.t. \, ||\mathbf{y} - \mathbf{A}\mathbf{x}||_2 \leq \epsilon$$

  J is non-smooth sparsity promoting function, e.g., ℓ_0-norm, ℓ_1-norm.
- Existing liteature has focussed on J being convex.
- This work focusses on J being nonconvex & non-smooth.
- Approach: use proximal algorithms.
- Proximal mapping of function g is defined as

$$prox_g(x) = \underset{u \in dom(g)}{\arg\min} \left\{ \frac{1}{2} ||x - u||_2^2 + g(u) \right\}$$

- Examples:
  - $g(x) = \lambda ||\mathbf{x}||_0$, $prox_g$ is the hard thresholding operator
  - $g(x) = \lambda ||\mathbf{x}||_1$, $prox_g$ is the soft thresholding operator

Sparse Signal Recovery Using Iterative Proximal Projection

- **Splitting methods**

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + g(\mathbf{x})$$

f is smooth, convex/nonconvex and g is non-smooth, nonconvex

- Forward-Backward Splitting

$$\mathbf{x}_{k+1} = prox_{\mu_k g}\left(\mathbf{x}_k - \mu_k \nabla f(\mathbf{x}_k)\right)$$

- Backward-Backward Splitting

$$\mathbf{x}_{k+1} = prox_g\left(prox_{\mu_k f}(\mathbf{x}_k)\right)$$

Sparse Signal Recovery Using Iterative Proximal Projection

- **Splitting methods**

$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) + g(\mathbf{x})$$

f is smooth, convex/nonconvex and g is non-smooth, nonconvex

  - Forward-Backward Splitting

$$\mathbf{x}_{k+1} = prox_{\mu_k g}\left(\mathbf{x}_k - \mu_k \nabla f(\mathbf{x}_k)\right)$$

  - Backward-Backward Splitting

$$\mathbf{x}_{k+1} = prox_g\left(prox_{\mu_k f}(\mathbf{x}_k)\right)$$

- **Accelerations**

$$\begin{aligned}\hat{\mathbf{x}}_k &= \mathbf{x}_k + w(\mathbf{x}_k - \mathbf{x}_{k-1}) \\ \mathbf{x}_{k+1} &= prox_{\mu_k g}\left(\hat{\mathbf{x}}_k - \mu_k \nabla f(\hat{\mathbf{x}}_k)\right)\end{aligned}$$

Learning Convex Regularizers for Optimal Bayesian Denoising

Authors: Ha Q. Nguyen, Emrah Bostan and Michael Unser

- Recover stochastic x from its noisy observations

$$y = x + n$$

  - n is AWGN of variance σ^2.
  - MAP inference is the way to solve this
  - Revisit MAP from perspective of estimation accuracy instead of deviation from prior model

Learning Convex Regularizers for Optimal Bayesian Denoising

Authors: Ha Q. Nguyen, Emrah Bostan and Michael Unser

- Recover stochastic x from its noisy observations

$$y = x + n$$

  - n is AWGN of variance σ^2.
  - MAP inference is the way to solve this
  - Revisit MAP from perspective of estimation accuracy instead of deviation from prior model
- Typical MAP formulation:

$$\hat{\boldsymbol{x}}_{\text{MAP}} = \arg\min_{\boldsymbol{x}} \left\{ \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{x}\|_2^2 + \sigma^2 \sum_{i=1}^{N} \Phi_U([\boldsymbol{L}\boldsymbol{x}]_i) \right\},$$

  - L = Whitening filter
  - $\Phi_U = -log\ p_U$ is called the penalty function
- The penalty function is designed such that it captures the statistics of collection of clean signals.

Learning Convex Regularizers for Optimal Bayesian Denoising

Authors: Ha Q. Nguyen, Emrah Bostan and Michael Unser

- Recover stochastic x from its noisy observations

$$y = x + n$$

- ADMM based denoising solution is proposed.
- Remarks:
  - Assumed that $\mathbf{x}$ can be whitened by some matrix $\mathbf{L}$.
  - $\mathbf{u} = \mathbf{Lx}$ has i.i.d. entries.
  - Penalty function Φ_U is separable.
  - ADMM formulation of MAP:

$$\frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2 + \sigma^2 \Phi_U(\boldsymbol{u}) - \langle \alpha, \boldsymbol{Lx} - \boldsymbol{u} \rangle + \frac{\mu}{2}\|\boldsymbol{Lx} - \boldsymbol{u}\|_2^2$$

Learning Convex Regularizers for Optimal Bayesian Denoising

Authors: Ha Q. Nguyen, Emrah Bostan and Michael Unser

- ADMM formulation of MAP:

$$\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{x}\|_2^2+\sigma^2\Phi_U(\boldsymbol{u})-\langle\alpha,\boldsymbol{L}\boldsymbol{x}-\boldsymbol{u}\rangle+\frac{\mu}{2}\|\boldsymbol{L}\boldsymbol{x}-\boldsymbol{u}\|_2^2$$

- Update for **u** looks like:

$$\boldsymbol{u}^{(k+1)}=\mathrm{prox}_{\sigma^2/\mu\Phi_U}\left(\boldsymbol{L}\boldsymbol{x}^{(k+1)}-\frac{1}{\mu}\alpha^{(k+1)}\right).$$

  - Update for **u** is typically a pointwise shrinkage operation.
  - Shrinkage operator depends on the choice of penalty Φ_U.
  - **Key Idea:** Learn optimal shrinkage directly from the data.

Learning Convex Regularizers for Optimal Bayesian Denoising

Authors: Ha Q. Nguyen, Emrah Bostan and Michael Unser

- Learning the shrinkage function from data:
  - Parameterized model for shrinkage function $T : \mathbb{R} \to \mathbb{R}$.

$$T(x) = \sum_{m=-M}^{M} c_m \psi\left(\frac{x}{\Delta} - m\right).$$

  - Given a collection of ground-truth signals $\{\mathbf{x}_l\}_{l=1}^{L}$, the parameters of shrinkage function can be learned by minimizing:

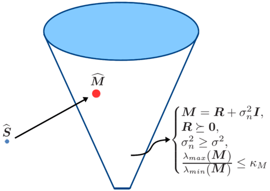
$$J(\boldsymbol{c}) = \frac{1}{2}\sum_{\ell=1}^{L}\left\|\boldsymbol{x}^{(K)}(\boldsymbol{c}, \boldsymbol{y}_\ell) - \boldsymbol{x}_\ell\right\|_2^2.$$

  - Shrinkage operator learned at one noise level works for all noise levels!

A Geometric Approach to Covariance Matrix Estimation and Applications to Radar Problems

Authors: Augusto Aubrey, Antonio De Maio and Luca Pallotta

- Given data vectors $\mathbf{r}_1, \mathbf{r}_{2,K}$, estimate the underlying true covariance matrix subject to certain constraints.
  - Step-1: First compute the sample covariance matrix $\hat{\mathbf{S}} = \frac{1}{K}\sum_{i=1}^{K} \mathbf{r}_i \mathbf{r}_i^T$.
  - Step-2: Project $\hat{\mathbf{S}}$ into a specific set in some matrix norm sense (unitary invariant).

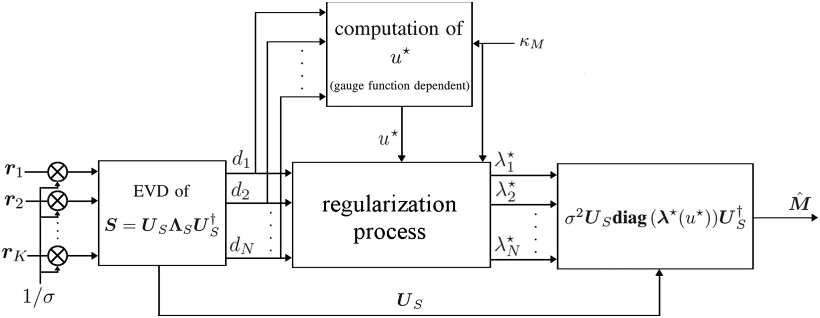


- The constraints encompasses p.d. matrices which can be modeled as sum of an unknown psd matrix(interference + clutter) and a term proportional to Identity matrix (white noise).

A Geometric Approach to Covariance Matrix Estimation and Applications to Radar Problems

Authors: Augusto Aubrey, Antonio De Maio and Luca Pallotta

- Proposed covariance matrix estimation flow



Other Interesting Papers:

- Optimized Self-Localization for SLAM in Dynamic Scenes Using Probability Hypothesis Density Filters
- Stochastic Approximation and Memory-Limited Subspace Tracking for Poisson Streaming Data
- Optimal Nested Test Plan for Combinatorial Quantitative Group Testing
- On Fienup Methods for Sparse Phase Retrieval
- Complex Factor Analysis and Extensions
- Nesterov-Based Alternating Optimization for Nonnegative Tensor Factorization: Algorithm and Parallel Implementation